Fully Modified Least Squares Estimation and Vector Autoregression of Models with Seasonally Integrated Processes

by

Gilbert Chiu-sing Lui

Thesis submitted in June 1997

in partial fulfilment of the requirements

for the degree of

Master of Philosophy in Economics

Department of Economics

The Chinese University of Hong Kong

Shatin, Hong Kong

September 25, 1997

Acknowledgment

I would like to thank my thesis advicer, Professor Terence Tai-leung Chong, who gives me invaluable comments. This paper is dedicated to Betty Chang.

Abstract

In this paper, we study the fully modified (FM) estimation of models with an unknown mixture of stationary and seasonally integrated series. As the FM-OLS estimator suggested by Phillips (1995) is incapable of estimating our model, we propose a more general formula of FM estimator. We denote this estimator by FM-SEA estimator. The $I(0)$ component has a normal limit distribution with different asymptotic variance from that derived by Phillips (1995). On the other hand, the limiting distribution of nonstationary component is mixed normal and proportional to the result of Phillips (1995, Theorem 4.1 (b)) by the order of seasonal integration. We also introduce a deterministic trend component into our model. The asymptotic distributions of stochastic trend and deterministic trend components are mixed. This mixture of the variates is different from that of Phillips (1995).

We also compare the finite sample properties of FM-SEA and OLS estimators through Monte Carlo experiments. The main result is that the average bias and RMSE of FM-SEA estimator are smaller than those of OLS estimator for stochastic trend component as the sample size is increased. Also, the t-statistics for FM-SEA estimators have smaller average bias and root mean squared error especially when the sample size is enlarged. Thus, the FM-SEA estimator works better than OLS estimator as far as hypothesis testing is concerned.

Table of Contents

List of Tables

List of Figures

1. Introduction

The purpose of this paper is to estimate the models with an unknown number of seasonally integrated processes. The complexity of studying seasonally integrated process is arisen from the presence of both real unit roots and complex unit roots. For the models with $I(1)$ processes, an extensive literature exists for the limit theories of this type of models (e.g. see Dickey and Fuller (1979), Johansen (1988, 1991) and Phillips (1986, 1987, 1988)). The OLS coefficient estimate of an I(1) variable is $O(T)$-consistent but not asymptotically normal. Park and Phillips (1988), Sims et al (1990) and West (1988) showed that the limiting distribution of OLS coefficient estimate of an $I(1)$ variable can be normal in some models. The results in the aforementioned papers require the prior knowledge of data, that is, unit root tests should be performed before estimation. To solve the inference problem arising from the non-standard limit distribution of integrated processes, Hansen and Phillips (1990) developed the fully modified least squares method. The distribution of FM-OLS coefficient estimates is asymptotically normal. Phillips (1995) (denoted PFM) extended the results by including an unknown number of unit roots. By this method, no unit root tests are required before estimation.

In the past decades, some researchers have worked on the time series which exhibit considerable seasonality. For example, Davidson et al (1978) estimated the consumption function in ECM form with the quarterly seasonally unadjusted UK consumption expenditures. Osborn et al (1988) examined the seasonality of UK consumption data by using the diagnostic tests developed by Dickey and Fuller (1979), Dickey et al (1984), Hylleberg et al (1990) (denoted HEGY) and Hasza and Fuller (1982). Concerning with the theoretical works, Dickey et al (1984) derived the limiting distribution of OLS coefficient estimate of seasonally integrated variable. The OLS estimate is $O_p(T-1)$ and has a non-standard limit distribution. Recently, HEGY and Lee (1992) developed the seasonal cointegration tests. Their tests are based on the OLS estimation of an ECM model and Johansen (1988, 1991)'s maximum likelihood approach respectively. Similar to the models with $I(1)$ process, the OLS estimation of these models requires the seasonal unit root tests beforehand.

In this paper, we incorporate the methodology of PFM to estimate seasonally integrated process. The paper is organised as follows. Section 2 describes the model and states the assumptions. Then, the limit theory of OLS estimate of seasonally integrated component is derived by using functional central limit theorem as nobody has derived the result rigorously before. In Section 3, we propose a more general form of FM estimator for a seasonally integrated model and derive the limiting distributions of the parameter estimates. In Section 4, a VAR model in error correction form will be estimated by using the theorem derived in Section 3. In particular, a VAR model with quarterly data will be presented afterwards. Monte Carlo experiments which compare the finite sample properties of FM-SEA and OLS estimators under various orders of seasonal integration and various degrees of contemporaneous correlation of errors will be performed in Section 5. Finally, the conclusion is drawn in Section 6.

The notations used in this paper are followed from the standard time series econometrics and PFM. Thus, we denote the long-run variance and one-sided long-run covariance matrices of the stationary time series u_t as $\Omega = \sum_{j=-\infty}^{\infty} E(u_1 u_{1-j}')$ and $\Delta = \sum_{j=0}^{\infty} E(u_1 u_{1-j}')$ respectively. We write the vector Brownian motion with mean zero and covariance matrix Ω as BM(Ω) and the integrals with respect to the Lebesgue measure, like $\int_0^1 B(r)dr$, as $\int_0^1 B$. The notation $y_t \equiv I(d)$ signifies the time series y_t to be integrated of order d, where d is a non-negative integer, and the long-run variance matrix of $\Delta^d y_t = (1-L)^d y_t$ is positive definite. Similarly, we denote $\Delta_d y_t = (1-L^d)y_t = y_t - y_{t-d}$ which also satisfies the same condition as $\Delta^d y_t$ does. Furthermore, $y_t = SI(d,\theta)$ signifies the time series y_t to be seasonally integrated of order d at frequency $\theta \in [0,\pi)$. We define the inequality ">0" to be the positive definiteness of the matrices. We use the symbols "$\rightarrow_p$", "$\rightarrow_d$" and "$\equiv$" to signify the convergence in probability, convergence in distribution and equality in distribution respectively. Also, we use $\|A\|$ to signify the usual Euclidean norm $(tr(A'A))^{1/2}$, $|A|$ to denote the determinant of A, $vec(A)$ to stack the rows of matrix A into a column vector, $[x]$ to denote the integral part of the real finite number x. Finally, all limits in this paper are taken as the sample size $T \to \infty$ unless otherwise specified.

2. Models and Assumptions

Consider a linear model

$$y_t = Ax_t + u_{0,t} \qquad for\ t = 1, 2, ..., T, \tag{2.1}$$

where the observations at $t = 0, -1, -2, ..., -d+1$ are initialised by some given distributions, y_t and x_t are n-element and $m = (m_1+m_2)$-element column vectors respectively, coefficient matrix A is of order $(n \times m)$ and $T = \kappa d$ where κ is a positive integer. The parameter κ can be regarded as the number of years. A $(m \times m)$ orthogonal matrix $D = [D_1, D_2]$, satisfying condition $D'D = DD' = I$, is used to rotate the regressor space of (2.1). The regressor space is then specified according to

$$D_1'x_t = x_{1,t} = u_{1,t} \text{ and } D_2'\Delta_d x_t = \Delta_d x_{2,t} = u_{2,t},$$

where $d \geq 1$. Note that the above seasonal component is stochastic rather than deterministic, because Canova and Hansen (1995) showed that many seasonal patterns are not constant over time. Model (2.1) can then be rewritten as

$$y_t = A_1 x_{1,t} + A_2 x_{2,t} + u_{0,t} \qquad for\ t = 1, 2, ..., T, \tag{2.2}$$

where $A_1 = AD_1$ and $A_2 = AD_2$. Obviously, $x_{1,t}$ and $x_{2,t}$ are $I(0)$ and $SI(d, \theta)$ processes respectively by this transformation. The orthogonal matrix D over here has the similar effect as that in Sims et al (1990). In their paper, the deterministic trends with different degrees and stochastic trends with different orders of integration in a VAR(1) system are separated. This transformation requires the prior knowledge of data, that is, unit root tests and cointegration tests should be performed before estimation. However, in the FM-SEA framework, we can estimate model (2.1) without the prior knowledge of D. Data matrices in model (2.2) are denoted as uppercase letters and then model (2.2) can be rewritten as

$$Y' = A_1 X_1' + A_2 X_2' + U_0', \tag{2.3}$$

with $X_1 = U_1$, $\Delta_d X_2 = U_2$ where $Y' = [y_1, y_2, ..., y_T]$, $X_1' = [x_{1,1}, x_{1,2}, ..., x_{1,T}]$, $X_2' = [x_{2,1}, x_{2,2}, ..., x_{2,T}]$, $U_0' = [u_{0,1}, u_{0,2}, ..., u_{0,T}]$, $U_1' = [u_{1,1}, u_{1,2}, ..., u_{1,T}]$ and $U_2' = [u_{2,1}, u_{2,2}, ..., u_{2,T}]$.

Let $u_t = \left(u_{0,t}', u_{1,t}', u_{2,t}'\right)'$ be a $(n+m)$-element column vector and $\varphi_t = u_{0,t} \otimes u_{1,t}$. Then, we assume that u_t satisfies the following conditions :

ASSUMPTION 2.1.

(a) $u_t = C(L)\varepsilon_t = \sum_{j=0}^{\infty} C_j \varepsilon_{t-j}$, $\sum_{j=0}^{\infty} j^a \|C_j\| < \infty$ for some constant $a > 1$, $|C(1)| \neq 0$;

(b) ε_t is i.i.d. with zero mean, variance matrix $\Sigma_{\varepsilon\varepsilon} > 0$ and finite fourth order cumulants;

(c) $E(\varphi_{t,j}) = E(u_{0,t+j} \otimes u_{1,t}) = 0$ for all $j \geq 0$.

Assumption 2.1 (a) shows that u_t is generated by a stationary MA(∞) process. Assumption 2.1 (c) shows that the subvectors of u_t, $u_{0,t}$ and $u_{1,t}$, are serially uncorrelated at all lagged periods. Assumption 2.1 (b) ensures the validity of functional central limit theorems for u_t and $u_t u_t'$.

The limiting processes of partial sums of φ_t and u_t are obtained by the multivariate extension of Phillips and Solo (1992, Theorem 3.4). Thus,

$$T^{-1/2} \sum_{i=0}^{d-1} \sum_{j=0}^{\kappa-1} \varphi_{T-(i+dj)} = T^{-1/2} \sum_{t=1}^{T} \varphi_t \rightarrow_d N(0, \Omega_{\varphi\varphi}), \tag{2.4}$$

where $\Omega_{\varphi\varphi} = \sum_{j=-\infty}^{\infty} E(u_{0,t} u_{0,t+j}' \otimes u_{1,t} u_{1,t+j}')$;

$$T^{-1/2} \sum_{i=0}^{d-1} \sum_{j=0}^{[([Tr]-1)/d]} u_{[Tr]-(i+dj)} = T^{-1/2} \sum_{t=1}^{[Tr]} u_t \rightarrow_d B(r) \equiv BM(\Omega), \tag{2.5}$$

with $\Omega = C(1)\Sigma_{\varepsilon\varepsilon}C(1)'$ for all $r \in [0,1]$.

The long-run variance matrix Ω is given by

$$
\begin{aligned}
\Omega &= \lim_{T\to\infty} E\left(T^{-1}\sum_{i=0}^{d-1}\sum_{j=0}^{\kappa-1} u_{T-(i+dj)} \sum_{i=0}^{d-1}\sum_{j=0}^{\kappa-1} u'_{T-(i+dj)}\right) \\
&= \lim_{T\to\infty} E\left(T^{-1}\sum_{t=1}^{T} u_t \sum_{t=1}^{T} u'_t\right) \\
&= \lim_{T\to\infty} \left[E\left(T^{-1}\sum_{t=1}^{T} u_t u'_t\right) + E\left(T^{-1}\sum_{t=1}^{T}\sum_{s=t+1}^{T} u_s u'_t\right)\right. \\
&\qquad \left. + E\left(T^{-1}\sum_{t=1}^{T}\sum_{s=t+1}^{T} u_t u'_s\right)\right] \\
&= E(u_t u'_t) + \left[\sum_{s=1}^{\infty} E(u_1 u'_{1-s}) + E(u_1 u'_{1+s})\right] \\
&= \Sigma + \Lambda + \Lambda',
\end{aligned}
$$

where $\Sigma = E(u_t u'_t)$, $\Lambda = \sum_{j=1}^{\infty} E(u_1 u'_{1-j})$. The one-sided long-run covariance matrix of u_t is defined by $\Delta = \Sigma + \Lambda = \sum_{j=0}^{\infty} E(u_1 u'_{1-j}) = \sum_{j=0}^{\infty} \Gamma(j)$ where $\Gamma(j)$ is the autocovariance matrix of u_t at the j^{th} period lag. The variance matrix Σ and long-run variance matrix Ω of u_t are partitioned conformably with the partitions of u_t and we denote the submatrices as Σ_{ij} and Ω_{ij} respectively for $i, j = 0, 1, 2$. Similarly, the Brownian motion B is partitioned conformably with the partitions of u_t, that is, $B = [B'_0, B'_1, B'_2]'$.

Now, we study the limiting process of seasonally integrated process, $x_{2,t}$. Note that when d is even,

$$1-L^d = (1-L)(1+L)(1+L^2+...+L^{d-2}) = (1-L)(1+L)\prod_{j=1}^{(d-2)/2}(1-2\cos\theta_j L+L^2);$$

when d is odd,

$$1-L^d = (1-L)(1+L^2+...+L^{d-1}) = (1-L)\prod_{j=1}^{(d-1)/2}(1-2\cos\theta_j L+L^2);$$

where $\theta_j = 2j\pi/d$. We can follow the componentwise analyses of Tiao and Tsay (1983, 1990), Chan and Wei (1988) and Chan (1989) to develop the limit theory of seasonally integrated process. However, Rao (1978) and Chan and Wei (1988) mentioned that the closed form of limiting distribution of least squares estimates were difficult to derive. We express $x_{2,t}$ as a sum of partial sum processes.

$$(1+L^2+...+L^{d-1})x_{2,t} = \sum_{j=1}^{t} u_{2,j} + x_{2,0} + x_{2,-1} + ... + x_{2,-d+1}. \tag{2.6}$$

Also, we can write the term $x_{2,t}$ as

$$x_{2,t} = \sum_{j=0}^{[(t-1)/d]} u_{2,t-dj} + x_{2,t-d[(t-1)/d]-d}. \tag{2.7}$$

The partial sum for $x_{2,t}$ represents the sum of errors $u_{2,t-dj}$ at the corresponding season. We observe that, under Assumption 2.1, $x_{2,t}$ converges in distribution to

$$(T/d)^{-1/2}x_{2,[Tr]} \to_d B_{2,1}(r) \equiv BM(\Omega_{22,1}), \tag{2.8}$$

for $r \in [0,1]$ where $\Omega_{22,1}$ corresponds to the long-run variance of $u_{2,t}$ at the first season. We define the limiting processes of $x_{2,t}$ at the other lags in (2.6) by $B_{2,j}(r)$ with long-run variances $\Omega_{22,j}$ for $1 \leq j \leq d$. In fact, the $B_{2,j}(r)'s$ for all j are not necessarily independent. As it is readily seen from (2.6), the limiting processes defined above satisfy the following condition :

$$d^{-1/2} \sum_{j=1}^{d} B_{2,j}(r) = B_2(r). \tag{2.9}$$

Lemma 2.1 gives the limiting distribution of OLS estimate of $SI(d, \theta)$ process.

LEMMA 2.1. Under Assumption 2.1, the followings hold :

(a) $T^{-1} U_2' X_{2,-d} \rightarrow_d d^{-1} \int_0^1 \sum_{j=1}^d dB_j B_j' + \Lambda_d$,

(b) $T^{-1} X_{2,-d}' X_{2,-d} \rightarrow_d d^{-2} \int_0^1 \sum_{j=1}^d B_j B_j'$,

where $X_{2,-d}$ denotes the data matrices of $x_{2,t-d}$, $\Lambda_d = \sum_{j=d}^{\infty} \Gamma_{22}(j)$ and $\Gamma_{22}(j)$ denotes the autocovariance of $u_{2,t}$ at the j^{th} period lag. Part (a) can be proved by using the results of Phillips (1988b, Lemma 2.5 (c) and Theorem 2.6 (a)). The result of part (b) follows the standard proof of the sample moment of $I(1)$ process by replacing the term T^{-2} with $(T/d)^{-2}$.

Lemma 2.1 generalises the univariate results of Hasza and Fuller (1982, Lemma 3.1 (iv) and (v)) and Dickey and Fuller (1979, Theorem 1). Note that the sample moments of $SI(d, \theta)$ process converge to the integrals of sum of cross products of Brownian motions at all seasons. For a usual $I(1)$ process, the result of part (b) is reduced to $d = 1$. This result is quite intuitive because, from (2.9), the sum of Brownian motions at all seasons equals the Brownian motion for the whole year times $d^{1/2}$.

THEOREM 2.2. Suppose $\widehat{G}$ is the OLS estimator by regression of $x_{2,t}$ on $x_{2,t-d}$, the asymptotic distribution of $\widehat{G}$ is determined by

$$\widehat{G} - I = (U_2' X_{2,-d})(X_{2,-d}' X_{2,-d})^{-1}. \tag{2.10}$$

When the conditions of Lemma 2.1 are satisfied, we have

$$T(\widehat{G} - I) \rightarrow_d \left(d \int_0^1 \sum_{j=1}^{d} dB_{2,j} B_{2,j}' + d^2 \Lambda_d \right) \left(\int_0^1 \sum_{j=1}^{d} B_{2,j} B_{2,j}' \right)^{-1}. \tag{2.11}$$

Clearly, $\widehat{G}$ has a non-standard limit distribution and is of order $O_p(T^{-1})$. Also, its asymptotic variance increases with the order of seasonal integration. The term $d^2\Lambda_d$ indicates the horizontal shift of non-degenerating distribution as compared with the case of i.i.d. errors. When d increases, the size of movement increases by d^2 times with fixed Λ_d. In a univariate case, under the assumption of i.i.d. errors, the limiting distribution of $\widehat{G}$ is reduced to

$$T(\widehat{G} - 1) \rightarrow_d \frac{d}{2} \left(\sum_{j=1}^{d} B_{2,j}^2(1) - d \right) \left(\int_0^1 \sum_{j=1}^{d} B_{2,j}^2 \right)^{-1}. \tag{2.12}$$

By using the inifinite expansion of Brownian Motion by kernel method (See Kac

(1980, p.15)), we have

$$B_{2,j}(r) = \sum_{g=0}^{\infty} \frac{2\sqrt{2}Z_{j,g}}{(2g+1)\pi} sin\left[\left(g+\frac{1}{2}\right)\pi\right], \tag{2.13}$$

where $Z_{j,g}$ are n.i.d.(0,1). Then, the result of Hasza and Fuller (1982, Corollary 3.1) can be derived by the same procedures as Chan and Wei (1988, p.382) did.

Now, we obtain the kernel estimates of Ω and Δ by the following formula (See Andrews (1991, p.820)) :

$$\widehat{\Omega} = \sum_{j=-T+1}^{T-1} w(j/K)\widehat{\Gamma}(j) \text{ and } \widehat{\Delta} = \sum_{j=0}^{T-1} w(j/K)\widehat{\Gamma}(j), \tag{2.14}$$

where $w(.)$ is a kernel function and K is a bandwidth parameter. The sample autocovariance matrix is estimated by

$$\widehat{\Gamma}(j) = T^{-1} \sum_{t=j+1}^{T} \widehat{u}_t \widehat{u}'_{t-j},$$

where $\widehat{u}_t = \left(\widehat{u}'_{0,t}, u'_{1,t}, u'_{2,t}\right)'$ and $\widehat{u}_{0,t}$ is obtained by OLS estimation of (2.1).

Note that the kernel estimates in (2.14) do not contain the degree of freedom adjustment because the adjustment term will vanish asymptotically. The sums in (2.13) are truncated when the weight $w(j/K) = 0$ for $|j| \geq K$ or the weight $w(j/K)$ diminishes to a bounded limit as $|j| \to K$. Since the autocovariances with too long period lags do not provide important information to the kernel estimates of Ω and Δ, they can be ignored. These properties of the kernel function will be given in Assumption 2.2.

ASSUMPTION 2.2. The kernel function is defined by $w(.) : R \to [-1, 1]$, which is a twice continuously differentiable even function and satisfies the following conditions :

(a) $w(0) = 1, w'(0) = 0, w''(0) \neq 0$ where $w'(.)$ and $w''(.)$ denote the first and second derivatives of $w(.)$ respectively and either

(b) $w(x) = O(x^{-2})$, as $|x| \to 1$, or

(b') $w(x) = 0, |x| \geq 1$, with $\lim_{|x| \to 1} w(x)/ (1 - |x|)^2 =$ constant.

Obviously, Assumption 2.2 (b) and (b') indicate the assumptions of untruncated and truncated kernels respectively. More specifically, Assumption 2.2 (b) allows for the quadratic spectral kernel which is defined by

$$w(x) = \begin{cases} 1 & \text{when } x = 0, \\ \dfrac{25}{12\pi^2 x^2} \left(\dfrac{\sin(6\pi x/5)}{6\pi x/5} - \cos(6\pi x/5) \right) & \text{when } x \neq 0, \end{cases}$$

while Assumption 2.2 (b') allows for the Parzen and Tukey-Hanning kernels which are defined by

Parzen Kernel :

$$w(x) = \begin{cases} 1 - 6x^2 + 6|x|^3 & \text{when } 0 \leq |x| \leq 1/2, \\ 2(1 - |x|)^3 & \text{when } 1/2 \leq |x| \leq 1, \\ 0 & \text{otherwise.} \end{cases}$$

Tukey-Hanning Kernel :

$$w(x) = \begin{cases} (1+\cos(\pi x))/2 & \text{when } |x| \leq 1, \\ 0 & \text{when } |x| > 1. \end{cases}$$

The conditions of kernel function stated above are sufficient for our proofs but not necessary. So, we may choose the other classes of kernel functions but more complicated expressions will be obtained in our proofs.

PFM showed that the growth of bandwidth parameter K affected the consistency of FM-OLS estimators, thus we assume that the bandwidth parameter K expands with T for our FM-SEA estimators.

ASSUMPTION 2.3. The bandwidth parameter K of the kernel function in the kernel estimates of Ω and Σ grows according to the following conditions :

(a) $K = O_e(T^k)$ for some $k \in (1/4, 2/3)$,

(b) $K = O_e(T^k)$ for some $k \in (1/4, 1)$,

(c) $K = O_e(T^k)$ for some $k \in (0, 2/3)$,

(d) $K = O_e(T^k)$ for some $k \in (0, 1)$,

where the order of expansion rate of K is defined as $K = O_e(T^k)$ if $K \sim c_T T^k$ when $T \to \infty$, where c_T is slowly varying at infinity, that is, $\lim_{T\to\infty} c_{Tx}/c_T = 1$ for any constant $x > 0$. An obvious example is given by $c_T = \ln T$.

Then, we have

$$c_{Tx}/c_T = (\ln T + \ln x)/\ln T = 1 + \ln x/\ln T \to 1.$$

Condition (a) implies that $K/T^{2/3} + T^{1/4}/K \to 0$ as $T \to \infty$. Note that conditions (a) and (b) rule out the case of optimal expansion rate with $k = 1/5$ at which the asymptotic truncated mean squared error of kernel estimates is minimised (See Andrews (1991)). The condition for optimal expansion rate is not necessary. In our analysis, we choose the expansion rate k for the convergence of estimated coefficients. However, we do not need to choose the growth rate of K in the empirical works. By using Assumption 2.2, the consistency of kernel estimate $\widehat{\Omega}$ is shown in the following :

$$\begin{aligned}
\widehat{\Omega} &= \sum_{j=-T+1}^{T-1} [w(j/K) - w(0)]\,\widehat{\Gamma}(j) + \sum_{j=-T+1}^{T-1} \widehat{\Gamma}(j) \\
&= K^{-1} \sum_{j=-T+1}^{T-1} w'(\omega_j)\widehat{\Gamma}(j) + \sum_{j=-T+1}^{T-1} \widehat{\Gamma}(j) \qquad \text{where } \omega_j \in (0, j/K)\ \forall j \\
&= \sum_{j=-T+1}^{T-1} \widehat{\Gamma}(j) + O_p(T^{-1/2}K^{-1}) \\
&\to_p \sum_{j=-\infty}^{\infty} \Gamma(j) = \Omega
\end{aligned}$$

Similarly, the consistency of kernel estimate $\widehat{\Delta}$ can also be followed directly. Note that the consistency of $\widehat{\Omega}$ and $\widehat{\Delta}$ holds for $k \in (0,1)$. In PFM's framework, a consistent kernel estimate of the long-run variance matrix of disturbance terms is used to correct the OLS estimator. This kernel estimate of the long-run variance matrix can help the convergence of both $I(0)$ and $I(1)$ components when the bandwidth parameter satisfies the condition $k \in (1/4, 2/3)$. Without pretesting the data, the distributions of FM-OLS estimators for $I(0)$ and $I(1)$ components were asymptotically normal and mixed normal respectively. Also, the optimality of FM-OLS estimator for $I(1)$ component was shown by Phillips (1991a, equation (7)). We will show the results for seasonally integrated processes in the next section.

3. Asymptotics of FM-SEA Estimators

3.1. Model without Deterministic Trend

Before we propose the fully modified estimator, let us modify the OLS estimator of A by three different directions. First, the variable y_t in model (2.1) is transformed by

$$y_t^+ = y_t - \widehat{\Omega}_{0x}\widehat{\Omega}_{xx}^{-1}\Delta_d x_t, \tag{3.1}$$

where $\widehat{\Omega}_{0x}$ and $\widehat{\Omega}_{xx}$ are the kernel estimates of long-run covariance matrices of $(\widehat{u}_{0,t}, \Delta_d x_t)$ and $(\Delta_d x_t, \Delta_d x_t)$ respectively. According to the above equation, $\widehat{\Omega}_{0x}\widehat{\Omega}_{xx}^{-1}\Delta_d x_t$ is the projected values of $u_{0,t}$ on the space of $\Delta_d x_t$. The endogeneity arising from the contemporaneous correlation between $u_{0,t}$ and $u_{2,t}$ is then eliminated by (3.1), so that the variable y_t^+ is independent of $\Delta_d x_{2,t} = u_{2,t}$. According to Park and Phillips (1988), cointegrating regression yields the consistent and asymptotically mixed normal OLS coefficient estimates with the independence of $u_{0,t}$ and $u_{2,t}$. Note that the $x_{1,t}$ component of x_t is over-differenced by this transformation. The problem of over-differencing yields a singular matrix of Ω_{xx}. Nevertheless, the kernel estimate $\widehat{\Omega}_{0x}\widehat{\Omega}_{xx}^{-1}$ eliminates the singularity problem at some bandwidth. This point has been discussed by PFM (p.1059) and we will illustrate the point later in this section.

Second, the serial correlation correction term is given by

$$\widehat{\Delta}_{0x}^{+} = \widehat{\Delta}_{0x} - \widehat{\Omega}_{0x}\widehat{\Omega}_{xx}^{-1}\widehat{\Delta}_{xx}, \tag{3.2}$$

where $\widehat{\Delta}_{0x}$ and $\widehat{\Delta}_{xx}$ are the kernel estimates of one-sided long-run covariance matrices of $(\widehat{u}_{0,t}, \Delta_d x_t)$ and $(\Delta_d x_t, \Delta_d x_t)$ respectively. The transformation on $\widehat{\Delta}_{0x}$ removes the serial covariance of regression error $u_{0,t}$ and lagged values $u_{2,t-j}$ for $j > 0$. This is because the past persistence effect of $u_{2,t}$ in $x_{2,t}$ induces the one-sided long-run covariance matrix $\widehat{\Delta}_{0x}$ to carry this second order bias during OLS estimation of (2.1).

Third, we choose $S(X, d)$ to be the instrumental variable where $S(X, d) = \sum_{j=0}^{d-1} X_{-j}$ and X_{-j} denotes the data matrix of x_{t-j}. This instrumental variable serves two purposes. One is to correct the inconsistency of stationary component and the other is to make the asymptotic distribution of non-stationary component to have a simpler functional form as shown in the coming up theorem.

In (2.3), we have decomposed the regressor x_t into $x_{1,t}$ and $x_{2,t}$. However, we do not want to estimate A_1 and A_2 directly. This is because this estimation requires the prior knowledge of D. Also, we concentrate on the estimation of A, in model (2.1), for which no prior knowledge of transformation matrix D is required. Then, combining (3.1), (3.2) and the instrumental variables, the FM-SEA estimation of coefficient matrix A in model (2.1) is given by

$$\widehat{A}^{+} = \left(Y^{+\prime}S(X,d) - T\widehat{\Delta}_{0x}^{+}\right)\left(X'S(X,d)\right)^{-1}. \tag{3.3}$$

The matrices $\widehat{A}_1^{+}$ and $\widehat{A}_2^{+}$ are obtained by $\widehat{A}_1^{+} = \widehat{A}^{+}D_1$ and $\widehat{A}_2^{+} = \widehat{A}^{+}D_2$ respectively. Our proposed formula can allow for a wider class of models. When $d = 1$, this formula is reduced to the one by PFM. We interpret (3.3) by rewriting it as

$$\begin{aligned}\widehat{A}^{+} &= \left(Y'S(X,d) - T\widehat{\Delta}_{0x}\right)\left(X'S(X,d)\right)^{-1} \\ &\quad -\widehat{\Omega}_{0x}\widehat{\Omega}_{xx}^{-1}\left(\Delta_d X'S(X,d) - T\widehat{\Delta}_{xx}\right)\left(X'S(X,d)\right)^{-1}.\end{aligned} \tag{3.4}$$

The first term is the bias-corrected estimator suggested by Phillips and Hansen (1990) with $d = 1$ where all regressors are generated by $I(1)$ processes. This bias-corrected estimator is derived by using the results of Park and Phillips (1988, Theorem 3.1). Note that a scalar T is attached to the bias-correction term $\widehat{\Delta}_{0x}^{+}$. The reason is provided as follows. When x_t consists of $I(1)$ series, $U_0'X$ and $X'X$ are of order $O_p(T)$ and $O_p(T^2)$ respectively. The term $\left(\Delta_d X'S(X,d) - T\widehat{\Delta}_{xx}\right)(X'S(X,d))^{-1}$ is the bias-corrected estimator by instrumental variable regression of $\Delta_d x_t$ on x_t. Therefore, $\widehat{A}^+$ in (3.4) modifies the bias-corrected estimator of Phillips and Hansen (1990) by a projection on the bias-corrected estimator of $\Delta_d x_t$ on x_t with $d=1$, and we further introduce the instrumental variable $S(X,d)$ to adapt to the seasonally integrated processes.

Now, we go back to show why over-differencing $x_{1,t}$ does not affect the estimation result of $I(0)$ component asymptotically. The limiting distribution of $I(0)$ component of $\widehat{A}^+$ is determined by the term

$$
\begin{aligned}
& T^{1/2}\left[\left(T^{-1}U_0'S(X_1,d) - \widehat{\Delta}_{0\Delta_d u_1}\right) - \widehat{\Omega}_{0b}\widehat{\Omega}_{bb}^{-1}\left(T^{-1}U_b'S(X_1,d) - \widehat{\Delta}_{b\Delta_d u_1}\right)\right] \\
= \; & T^{1/2}\left(T^{-1}U_0'S(X_1,d) - \widehat{\Delta}_{0\Delta_d u_1}\right) + O_p(T^{1/2}K^{-2}) + O_p(K^{-1/2}) \\
& + O_p(T^{-1/2}K^{1/2}) \\
= \; & T^{-1/2}U_0'X_1 + o_p(1),
\end{aligned}
$$

where $k \in (0,1/4)$, see the proof of Theorem 3.1 in the appendix. The last two lines are obtained by Lemma 7.5 (b) and (c). Once the condition $k \in (0,1/4)$ has satisfied, the projected value of $\widehat{u}_{0,t}$ on the space $\Delta_d x_{1,t}$ vanishes asymptotically. Otherwise, it leads to the inconsistency of $\widehat{A}_1^+$. Further, the instrumental variable $S(X,d)$ for $I(0)$ component eliminates the corresponding one-sided long-run covariance matrix $\widehat{\Delta}_{0\Delta_d u_1}$ asymptotically. Thus, the limiting distribution of $\widehat{A}_1^+$ is determined by a term as OLS does.

We use (3.3) to derive the following theorem.

THEOREM 3.1. Under Assumptions 2.1, 2.2 and 2.3,

(a) $T^{1/2}\left(\widehat{A}_1^+ - A_1\right)$

$$\to_d N\left(0, \left(I \otimes \left(\textstyle\sum_{j=0}^{d-1} \Gamma_{11}(j)\right)^{-1}\right) \Omega_{\varphi\varphi} \left(I \otimes \left(\textstyle\sum_{j=0}^{d-1} \Gamma_{11}'(j)\right)^{-1}\right)\right),$$

(b) $T\left(\widehat{A}_2^+ - A_2\right) \to_d d\left(\int_0^1 dB_{0.2}B_2'\right)\left(\int_0^1 B_2 B_2'\right)^{-1}$

$$\equiv MN\left(0, d^2\Omega_{00.2} \otimes \left(\textstyle\int_0^1 B_2 B_2'\right)^{-1}\right),$$

where $B_{0.2} = B_0 - \Omega_{02}\Omega_{22}^{-1}B_2 \equiv BM(\Omega_{00.2})$, $\Omega_{00.2} = \Omega_{00} - \Omega_{02}\Omega_{22}^{-1}\Omega_{20}$ and $\Gamma_{11}(j)$ represents the autocovariance of $u_{1,t}$ at the j^{th} period lag. Note that part (a) holds when Assumption 2.3 (b) is satisfied and part (b) holds when Assumption 2.3 (c) is satisfied. For both part (a) and (b) to hold, Assumption 2.3 (a) should be satisfied.

By Theorem 3.1, the FM-SEA estimators are consistent and their distributions are different from those of PFM. The distributions of coefficient estimates of $I(0)$ and $SI(d,\theta)$ components are asymptotically normal and mixed normal respectively. In particular, the limit distribution of $\widehat{A}_2^+$ equals d multiple of the result by Theorem 4.1 (b) of PFM. This implies that its asymptotic variance is d^2 times of PFM. Similar to the result of Hasza and Fuller (1982, Corollary 3.1), the non-standard limit distribution of least squares estimate also depends on the value of d. According to Theorem 2.2, the limiting distribution of non-stationary component is expressed in the stochastic integrals of sum of limiting processes $x_{2,t}$ at all seasons. Now, the instrumental variable $S(X,d)$ here serves to simplify the functional form of asymptotic distribution of $\widehat{A}_2^+$.

On the other hand, the asymptotic variance of $\widehat{A}_1^+$ can be regarded as the result of standard instrumental variable estimation. However, we cannot compare the variance with that of PFM directly because we should identify the $I(0)$ component beforehand and realise the signs of $\Gamma_{11}(j)$ for $j = 1, 2, ..., d-1$. Also, the FM-SEA estimator $\widehat{A}_1^+$ is not necessarily optimal because the error $u_{0,t}$ is serially correlated as shown in Assumption 2.1 (a). The FM-SEA estimator $\widehat{A}_2^+$, which corresponds to the non-stationary component, is optimal as shown in Phillips (1991) in which a cointegrated system is estimated by maximising the log-likelihood conditional on $u_{2,t}$ under Gaussian error assumption.

COROLLARY 3.2. When $m_1 = 0$ in model (2.1), that is, $D = D_2$, and Assumptions 2.1, 2.2 and 2.3(d) are satisfied, we have

$$T\left(\widehat{A}_2^+ - A_2\right) \to_d d\left(\int_0^1 dB_{0.2}B_2'\right)\left(\int_0^1 B_2B_2'\right)^{-1}.$$

This result is quite trivial because the regressor space contains the non-stationary component only. This case is similar to the model studied by Phillips and Hansen (1990) in which x_t is generated by $I(1)$ processes. Consequently, the limiting distribution of $\widehat{A}_2^+$ in this case can capture the whole bandwidth $k \in (0,1)$. Hence, Corollary 3.2 can be seen as the generalisation of the results of Phillips and Hansen (1990, Theorem 3.2).

COROLLARY 3.3. When $m_2 = 0$ in model (2.1), that is, $D = D_1$, and Assumptions 2.1, 2.2 and 2.3(b) are satisfied, we have

$$T^{1/2}\left(\widehat{A}_1^+ - A_1\right) \to_d N\left(0, \left(I \otimes \left(\sum_{j=0}^{d-1}\Gamma_{11}(j)\right)^{-1}\right)\Omega_{\varphi\varphi}\left(I \otimes \left(\sum_{j=0}^{d-1}\Gamma_{11}'(j)\right)^{-1}\right)\right).$$

In this corollary, the transformation of y_t and the use of the instrumental variable are not necessarily. If we estimate the model in this corollary by OLS, a simpler result can be obtained. Nevertheless, no prior knowledge is provided on the stationarity of the variable x_t. We may still apply the FM-SEA estimation to this model. Note that the bandwidth parameter K still grows at $k \in (1/4, 1)$ in this case, the reason has been provided in the appendix.

From Theorem 3.1, the limiting distribution of full coefficient matrix $\widehat{A}^+$ is given by

$$\begin{aligned} & T^{1/2}\left(\widehat{A}^+ - A\right) \\ = \;& T^{1/2}\left(\widehat{A}^+ - A\right)DD' \\ = \;& T^{12}\left(\widehat{A}_1^+ - A_1\right)D_1' + T^{1/2}\left(\widehat{A}_2^+ - A_2\right)D_2' \end{aligned}$$

$$\to_d \quad N\left(0, \left(I \otimes D_1 \left(\sum_{j=0}^{d-1} \Gamma_{11}(j)\right)^{-1}\right) \Omega_{\varphi\varphi} \left(I \otimes \left(\sum_{j=0}^{d-1} \Gamma'_{11}(j)\right)^{-1} D'_1\right)\right) \tag{3.5}$$

The $SI(d, \theta)$ and $I(0)$ components are of order $O_p(T^{-1})$ and $O_p(T^{-1/2})$ respectively, so the non-stationary component does not have any asymptotic effect on the limiting distribution of $\widehat{A}^+$. Although the asymptotic variance of $\widehat{A}^+$ involves the unknown transformation matrix D_1, it can be consistently estimated by the formula suggested by PFM (p.1037). Since $u_{1,t} = D'_1 x_t$, we write the asymptotic variance of $\widehat{A}^+$ as

$$\begin{aligned}
&\left(I \otimes D_1 \left(\sum_{j=0}^{d-1} \Gamma_{11}(j)\right)^{-1}\right) (I \otimes D'_1) \Omega_{\varphi_x \varphi_x} (I \otimes D_1) \\
&\times \left(I \otimes \left(\sum_{j=0}^{d-1} \Gamma'_{11}(j)\right)^{-1} D'_1\right) \\
= \; &\left(I \otimes D_1 \left(\sum_{j=0}^{d-1} \Gamma_{11}(j)\right)^{-1} D'_1\right) \Omega_{\varphi_x \varphi_x} \left(I \otimes D_1 \left(\sum_{j=0}^{d-1} \Gamma'_{11}(j)\right)^{-1} D'_1\right),
\end{aligned}$$

where $\Omega_{\varphi_x \varphi_x} = \sum_{j=-K+1}^{K-1} w(j/K) \widehat{\Gamma}_{\varphi_x \varphi_x}(j)$ and $\varphi_x = u_{0,t} \otimes x_t$. Together with $\left(T^{-1} X' S(X, d)\right)^{-1} = D_1 \left(T^{-1} D'_1 X' S(X, d) D_1\right)^{-1} D'_1 = D_1 \left(T^{-1} X'_1 S(X_1, d)\right)^{-1} D'_1$ $= D_1 \left(T^{-1} U'_1 S(U_1, d)\right)^{-1} D'_1 \to_p D_1 \left(\sum_{j=0}^{d-1} \Gamma_{11}(j)\right)^{-1} D'_1$, the asymptotic variance of $\widehat{A}^+$ is estimated by

$$\left(I \otimes \left(T^{-1}X'S(X,d)\right)^{-1}\right)\Omega_{\widehat{\varphi}_x\widehat{\varphi}_x}\left(I \otimes \left(T^{-1}\left(X'S(X,d)\right)'\right)^{-1}\right), \tag{3.6}$$

where $\Omega_{\widehat{\varphi}_x\widehat{\varphi}_x} = \sum_{j=-K+1}^{K-1} w(j/K)\widehat{\Gamma}_{\widehat{\varphi}_x\widehat{\varphi}_x}(j)$ and $\widehat{\varphi}_x = \widehat{u}_{0,t} \otimes x_t$. Suppose that we test the hypothesis $H_0 : Rvec(A) = r$, where R is of order $(q \times mn)$ and $rank(R)$ $= q$ and r is of order $(q \times 1)$. Then, the Wald statistic is given by

$$\begin{aligned} W^+ &= T(Rvec(A) - r)'\left[R\left(I \otimes \left(T^{-1}X'S(X,d)\right)^{-1}\right)\Omega_{\widehat{\varphi}_x\widehat{\varphi}_x}\right. \\ &\quad \times \left.\left(I \otimes \left(T^{-1}\left(X'S(X,d)\right)'\right)^{-1}\right)R'\right]^{-1}(Rvec(A) - r). \end{aligned} \tag{3.7}$$

Provided the condition

$$rank\left[R\left\{\left(I \otimes D_1\left(\sum_{j=0}^{d-1}\Gamma_{11}(j)\right)^{-1}\right)\Omega_{\varphi\varphi}\left(I \otimes \left(\sum_{j=0}^{d-1}\Gamma'_{11}(j)\right)^{-1} D'_1\right)\right\}R'\right] = q \tag{3.8}$$

holds, then

$$W^+ \to_d \chi^2_q.$$

The asymptotic normality of $\widehat{A}^+$ gives the standard result of hypothesis testing. No extra tabulation of critical values is required. The result holds when the asymptotic variance matrix of $R\widehat{A}^+$is non-singular.

EXAMPLE 3.1. Now, we illustrate the use of FM-SEA estimation by considering a special case of model (2.1) with $n = 1$ and $m = 2$. Explicitly,

$$y_t = a_1 x_{1,t} + a_2 x_{2,t} + u_{0,t}. \tag{3.9}$$

Now, x_t is decomposed to $x_{1,t}$ and $x_{2,t}$ which are $I(0)$ and $SI(d,\theta)$ variables respectively. The long-run covariance, variance and one-sided long-run covariance matrices of u_t are defined by

$$\Omega = \begin{bmatrix} \Omega_{00} & \Omega_{01} & \Omega_{02} \\ \Omega_{10} & \Omega_{11} & \Omega_{12} \\ \Omega_{20} & \Omega_{21} & \Omega_{22} \end{bmatrix}, \Sigma = \begin{bmatrix} \Sigma_{00} & \Sigma_{01} & \Sigma_{02} \\ \Sigma_{10} & \Sigma_{11} & \Sigma_{12} \\ \Sigma_{20} & \Sigma_{21} & \Sigma_{22} \end{bmatrix} \text{ and } \Delta = \begin{bmatrix} \Delta_{00} & \Delta_{01} & \Delta_{02} \\ \Delta_{10} & \Delta_{11} & \Delta_{12} \\ \Delta_{20} & \Delta_{21} & \Delta_{22} \end{bmatrix}.$$

Note that, in this single equation model, the submatrices of Ω, Σ and Δ are scalar. The variable y_t and one-sided long-run covariance matrix Δ_{0x} are transformed by (3.1) and (3.2).

$$\begin{aligned}
y_t^+ &= y_t - \begin{bmatrix} \widehat{\Omega}_{0\Delta_d x_1} & \widehat{\Omega}_{0\Delta_d x_2} \end{bmatrix} \begin{bmatrix} \widehat{\Omega}_{\Delta_d x_1 \Delta_d x_1} & \widehat{\Omega}_{\Delta_d x_1 \Delta_d x_2} \\ \widehat{\Omega}_{\Delta_d x_2 \Delta_d x_1} & \widehat{\Omega}_{\Delta_d x_2 \Delta_d x_2} \end{bmatrix}^{-1} \begin{bmatrix} \Delta_d x_{1,t} \\ \Delta_d x_{2,t} \end{bmatrix} \\
&= y_t - \begin{bmatrix} \widehat{\Omega}_{0\Delta_d u_1} & \widehat{\Omega}_{0u_2} \end{bmatrix} \begin{bmatrix} \widehat{\Omega}_{\Delta_d u_1 \Delta_d u_1} & \widehat{\Omega}_{\Delta_d u_1 u_2} \\ \widehat{\Omega}_{u_2 \Delta_d u_1} & \widehat{\Omega}_{u_2 u_2} \end{bmatrix}^{-1} \begin{bmatrix} \Delta_d u_{1,t} \\ u_{2,t} \end{bmatrix} \\
&= y_t - \frac{1}{\widehat{\Omega}_{\Delta_d u_1 \Delta_d u_1} \widehat{\Omega}_{u_2 u_2} - \widehat{\Omega}_{u_2 \Delta_d u_1} \widehat{\Omega}_{\Delta_d u_1 u_2}} \begin{bmatrix} \widehat{\Omega}_{0\Delta_d u_1} & \widehat{\Omega}_{0u_2} \end{bmatrix} \\
&\quad \times \begin{bmatrix} \widehat{\Omega}_{u_2 u_2} & -\widehat{\Omega}_{\Delta_d u_1 u_2} \\ -\widehat{\Omega}_{u_2 \Delta_d u_1} & \widehat{\Omega}_{\Delta_d u_1 \Delta_d u_1} \end{bmatrix} \begin{bmatrix} \Delta_d u_{1,t} \\ u_{2,t} \end{bmatrix} \\
&= y_t - \frac{1}{\widehat{\Omega}_{\Delta_d u_1 \Delta_d u_1} \widehat{\Omega}_{u_2 u_2} - \widehat{\Omega}_{u_2 \Delta_d u_1} \widehat{\Omega}_{\Delta_d u_1 u_2}}
\end{aligned}$$

$$
\begin{aligned}
&\times\left[\left(\widehat{\Omega}_{0\Delta_d u_1}\widehat{\Omega}_{u_2u_2}-\widehat{\Omega}_{0u_2}\widehat{\Omega}_{u_2\Delta_d u_1}\right)\Delta_d u_{1,t}\right.\\
&\left.+\left(\widehat{\Omega}_{0u_2}\widehat{\Omega}_{\Delta_d u_1\Delta_d u_1}-\widehat{\Omega}_{0\Delta_d u_1}\widehat{\Omega}_{\Delta_d u_1 u_2}\right)u_{2,t}\right]
\end{aligned}
$$

$$
\begin{aligned}
\widehat{\Delta}^+_{0x} &= \begin{bmatrix}\widehat{\Delta}^+_{0\Delta_d x_1} & \widehat{\Delta}^+_{0\Delta_d x_2}\end{bmatrix} = \begin{bmatrix}\widehat{\Delta}^+_{0\Delta_d x_1} & \widehat{\Delta}^+_{0u_2}\end{bmatrix}\\
&= \begin{bmatrix}\widehat{\Delta}_{0\Delta_d x_1} & \widehat{\Delta}_{0u_2}\end{bmatrix} - \begin{bmatrix}\widehat{\Omega}_{0\Delta_d u_1} & \widehat{\Omega}_{0u_2}\end{bmatrix}\begin{bmatrix}\widehat{\Omega}_{\Delta_d u_1\Delta_d u_1} & \widehat{\Omega}_{\Delta_d u_1 u_2}\\ \widehat{\Omega}_{u_2\Delta_d u_1} & \widehat{\Omega}_{u_2u_2}\end{bmatrix}^{-1}\\
&\quad\times\begin{bmatrix}\widehat{\Delta}_{\Delta_d u_1\Delta_d u_1} & \widehat{\Delta}_{\Delta_d u_1 u_2}\\ \widehat{\Delta}_{u_2\Delta_d u_1} & \widehat{\Delta}_{u_2u_2}\end{bmatrix}\begin{bmatrix}\widehat{\Delta}_{\Delta_d u_1\Delta_d u_1} & \widehat{\Delta}_{\Delta_d u_1 u_2}\\ \widehat{\Delta}_{u_2\Delta_d u_1} & \widehat{\Delta}_{u_2u_2}\end{bmatrix}\\
&= \begin{bmatrix}\widehat{\Delta}_{0\Delta_d x_1} & \widehat{\Delta}_{0u_2}\end{bmatrix} - \frac{1}{\widehat{\Omega}_{\Delta_d u_1\Delta_d u_1}\widehat{\Omega}_{u_2u_2}-\widehat{\Omega}_{u_2\Delta_d u_1}\widehat{\Omega}_{\Delta_d u_1 u_2}}\\
&\quad\times\begin{bmatrix}\widehat{\Omega}_{0\Delta_d u_1} & \widehat{\Omega}_{0u_2}\end{bmatrix}\begin{bmatrix}\widehat{\Omega}_{u_2u_2} & -\widehat{\Omega}_{\Delta_d u_1 u_2}\\ -\widehat{\Omega}_{u_2\Delta_d u_1} & \widehat{\Omega}_{\Delta_d u_1\Delta_d u_1}\end{bmatrix}\begin{bmatrix}\widehat{\Delta}_{\Delta_d u_1\Delta_d u_1} & \widehat{\Delta}_{\Delta_d u_1 u_2}\\ \widehat{\Delta}_{u_2\Delta_d u_1} & \widehat{\Delta}_{u_2u_2}\end{bmatrix}
\end{aligned}
$$

The submatrices of $\widehat{\Delta}^+_{0x}$ are

$$
\begin{aligned}
\widehat{\Delta}^+_{0\Delta_d x_1} &= \widehat{\Delta}_{0\Delta_d x_1} - \frac{1}{\widehat{\Omega}_{\Delta_d u_1\Delta_d u_1}\widehat{\Omega}_{u_2u_2}-\widehat{\Omega}_{u_2\Delta_d u_1}\widehat{\Omega}_{\Delta_d u_1 u_2}}\\
&\quad\times\left[\left(\widehat{\Omega}_{0\Delta_d u_1}\widehat{\Omega}_{u_2u_2}-\widehat{\Omega}_{0u_2}\widehat{\Omega}_{u_2\Delta_d u_1}\right)\widehat{\Delta}_{\Delta_d u_1\Delta_d u_1}\right.\\
&\quad\left.+\left(\widehat{\Omega}_{0u_2}\widehat{\Omega}_{\Delta_d u_1\Delta_d u_1}-\widehat{\Omega}_{0\Delta_d u_1}\widehat{\Omega}_{\Delta_d u_1 u_2}\right)\widehat{\Delta}_{u_2\Delta_d u_1}\right]\\
\widehat{\Delta}^+_{0u_2} &= \widehat{\Delta}_{0u_2} - \frac{1}{\widehat{\Omega}_{\Delta_d u_1\Delta_d u_1}\widehat{\Omega}_{u_2u_2}-\widehat{\Omega}_{u_2\Delta_d u_1}\widehat{\Omega}_{\Delta_d u_1 u_2}}\\
&\quad\times\left[\left(\widehat{\Omega}_{0\Delta_d u_1}\widehat{\Omega}_{u_2u_2}-\widehat{\Omega}_{0u_2}\widehat{\Omega}_{u_2\Delta_d u_1}\right)\widehat{\Delta}_{u_2\Delta_d u_1}\right.\\
&\quad\left.+\left(\widehat{\Omega}_{0u_2}\widehat{\Omega}_{\Delta_d u_1\Delta_d u_1}-\widehat{\Omega}_{0\Delta_d u_1}\widehat{\Omega}_{\Delta_d u_1 u_2}\right)\widehat{\Delta}_{u_2u_2}\right]
\end{aligned}
$$

Now, FM-SEA estimators $\widehat{a}_1^+$ and $\widehat{a}_2^+$ are obtained by (3.3)

$$\widehat{a}_1^+ = \frac{\begin{array}{c}(\sum_{t=1}^T S(x_{1,t},d)y_t^+ - T\widehat{\Delta}_{0\Delta_d x_1})(\sum_{t=1}^T S(x_{2,t},d)x_{2,t}) \\ -(\sum_{t=1}^T S(x_{2,t},d)y_t^+ - T\widehat{\Delta}_{0\Delta_d x_2})(\sum_{t=1}^T S(x_{1,t},d)x_{2,t})\end{array}}{\begin{array}{r}(\sum_{t=1}^T S(x_{1,t},d)x_{1,t})(\sum_{t=1}^T S(x_{2,t},d)x_{2,t}) - (\sum_{t=1}^T S(x_{2,t},d)x_{1,t}) \\ \times(\sum_{t=1}^T S(x_{1,t},d)x_{2,t})\end{array}} \tag{3.10}$$

$$\widehat{a}_2^+ = \frac{\begin{array}{c}(\sum_{t=1}^T S(x_{2,t},d)y_t^+ - T\widehat{\Delta}_{0\Delta_d x_2})(\sum_{t=1}^T S(x_{1,t},d)x_{1,t}) \\ -(\sum_{t=1}^T S(x_{1,t},d)y_t^+ - T\widehat{\Delta}_{0\Delta_d x_1})(\sum_{t=1}^T S(x_{2,t},d)x_{1,t})\end{array}}{\begin{array}{r}(\sum_{t=1}^T S(x_{1,t},d)x_{1,t})(\sum_{t=1}^T S(x_{2,t},d)x_{2,t}) - (\sum_{t=1}^T S(x_{2,t},d)x_{1,t}) \\ \times(\sum_{t=1}^T S(x_{1,t},d)x_{2,t})\end{array}} \tag{3.11}$$

The limiting distributions of $\widehat{a}_1^+$ and $\widehat{a}_2^+$ can be derived by Theorem 3.1. So, we have

COROLLARY 3.4. Under the conditions of Theorem 3.1 with $n = 1$, $m = 2$ and $d = 4$,

(a) $T^{1/2}\left(\widehat{a}_1^+ - a_1\right) \to_d N\left(0, \Omega_{\varphi\varphi}\left(\sum_{j=0}^{d-1}\Gamma_{11}(j)\right)^{-2}\right)$,

(b) $T\left(\widehat{a}_2^+ - a_2\right) \to_d d\left(\int_0^1 dB_{0.2}B_2\right)\left(\int_0^1 B_2^2\right)^{-1}$,

where $\Omega_{\varphi\varphi} = \sum_{j=-\infty}^{\infty} E(u_{0,t}u_{1,t}u_{0,t-j}u_{1,t-j})$, $B_{0.2} = B_0 - \Omega_{02}\Omega_{22}^{-1}B_2 \equiv BM(\Omega_{00.2})$, $\Omega_{00.2} = \Omega_{00} - \Omega_{02}\Omega_{22}^{-1}\Omega_{20}$. The conditions for this corollary to hold are the same as those in Theorem 3.1.

If $u_{0,t}$ and $u_{1,s}$ are i.i.d. and independent at all t and s, then $\Omega_{\varphi\varphi} = \Sigma_{00}\Sigma_{11}$ and the asymptotic variance of $\widehat{a}_1^+$ can be reduced to $\Sigma_{00}\Sigma_{11}\left(\sum_{j=0}^{d-1}\Gamma_{11}(j)\right)^{-2}$. Even under this restrictive assumption, we cannot obtain a simpler form of asymptotic variance. Only when $d = 1$ is the asymptotic variance reduced to the OLS result Σ_{00}/Σ_{11}. The result of part (b) is different from that of Hasza and Fuller (1982, Corollary 3.1) and Dickey et al (1984, Theorem 1). In those cases, the non-standard limiting distribution of $\widehat{a}_2^+$ is expressed by the ratio of weighted sum of chi-squared random variables with one degree of freedom.

3.2. Model with Deterministic Trends

PFM provided the limiting results of $\widehat{A}^+$ in the models with $I(0)$, $I(1)$ and deterministic trend components. Hansen (1992, Theorem 3) showed the asymptotics of FM-OLS estimator in single equation case with deterministic trend component but no intercept. These initiate us to generalise their results by introducing a deterministic trend component into the model. Now, we consider the model

$$y_t = Ax_t + \Pi p_t + u_{0,t} = \Phi z_t + u_{0,t}, \tag{3.12}$$

where Π and p_t are of order $(n \times q)$ and $(q \times 1)$ respectively. Suppose that the regressor x_t contains a determinsitic trend, then the variable x_t can be decomposed to

$$x_t = D_1 x_{1,t} + D_2 x_{2,t} + F p_t, \tag{3.13}$$

where F is of order $(m \times q)$. So, we can consider the transformation matrix D for z_t by $D = [D_L, D_M]$ where $D_L = \begin{bmatrix} D_1 \\ 0 \end{bmatrix}$ and $D_M = \begin{bmatrix} D_2 & F \\ 0 & I_q \end{bmatrix}$.

Then, (3.12) becomes

$$y_t = \Phi_1 z_{1,t} + \Phi_2 z_{2,t} + u_{0,t}, \tag{3.14}$$

where $\Phi_1 = \Phi D_L$, $\Phi_2 = \Phi D_M$, $z_{1,t} = x_{1,t}$ and $z_{2,t} = \left(x'_{2,t}, p'_t\right)'$. The variable p_t is defined by

$$p_t = (t^{s_1}, t^{s_2}, ..., t^{s_q})' \text{ with } 0 \leq s_1 < s_2 < ... < s_q,$$

for some finite integer s_i with $i = 1, 2, ..., q$. We generalise Hansen (1992) to include an intercept ($s_1 = 0$). The most common p_t's are the linear trend $(1, t)$ and quadratic trend $(1, t, t^2)$. Next, we define the limiting process of p_t by

$$\delta_T^{-1} p_{[Tr]} \to p(r) = (r^{s_1}, r^{s_2}, ..., r^{s_q})', \tag{3.15}$$

uniformly for all $r \in [0, 1]$ where $\delta_T = diag\,(T^{s_1}, T^{s_2}, ..., T^{s_q})$. The proof is given by Hansen (1992, p.90). Note that the limit functions $p(r)$ satisfy the condition $\int_0^1 pp' > 0$. For example, $p(r) = (1, r, r^2)'$, then

$$\int_0^1 pp' = \begin{pmatrix} 1 & 1/2 & 1/3 \\ 1/2 & 1/3 & 1/4 \\ 1/3 & 1/4 & 1/5 \end{pmatrix} > 0.$$

The FM-SEA regression on (3.12) yields the estimator

$$\widehat{\Phi}^+ = \left[\widehat{A}^+ \vdots \widehat{\Pi}^+\right] = \left(Y^{+\prime}S(Z,d) - \left[T\widehat{\Delta}^+_{0x} \vdots 0\right]\right)\left(Z'S(Z,d)\right)^{-1}, \tag{3.16}$$

where Z denotes the data matrix of z_t. Note that, in contrast to Theorem 3.2 and 3.3 of Park and Phillips (1988), the estimator $\widehat{\Pi}^+$ does not contain the bias-correction term. This is because the bias in deterministic trend component is arisen from the bias of non-stationary component. The removal of the latter bias eliminates the bias-correction term in the deterministic trend component automatically. The kernel estimate $\widehat{\Delta}^+_{0x}$ is obtained by using residuals $(\widehat{u}_{0,t}, \Delta_d \widehat{u}_{p,t})$; where the residual $\widehat{u}_{0,t} = y_t - \widehat{A}x_t - \widehat{\Pi}p_t$ is obtained by OLS regression of (3.12) and the residual $\widehat{u}_{p,t} = x_t - \widehat{F}p_t$ is obtained by OLS regression of x_t on p_t. Clearly, the OLS residuals are demeaned or detrended.

To show the appropriateness of using $\Delta_d \widehat{u}_{p,t}$ for the estimation of $\widehat{\Delta}^+_{0x}$, we consider the asymptotic properties of OLS estimate $\widehat{F}$. Specifically,

$$
\begin{aligned}
\widehat{F} &= (X'P)(P'P)^{-1} \\
&= \left[(D_1X_1' + D_2X_2' + FP')P\right](P'P)^{-1} \\
\widehat{F} - F &= D_1(X_1'P)(P'P)^{-1} + D_2(X_2'P)\left(P'P\right)^{-1} \\
T^{-1/2}\left(\widehat{F} - F\right)\delta_T &= T^{-1}D_1\left(X_1'P\left(T^{-1/2}\delta_T^{-1}\right)\right)^{-1}\left(T^{-1}\delta_T^{-1}P'P\delta_T^{-1}\right)^{-1} \\
&\quad + D_2\left(X_2'P\left(T^{-3/2}\delta_T^{-1}\right)\right)^{-1}\left(T^{-1}\delta_T^{-1}P'P\delta_T^{-1}\right)^{-1} \\
&= D_2\left(X_2'P\left(T^{-3/2}\delta_T^{-1}\right)\right)^{-1}\left(T^{-1}\delta_T^{-1}P'P\delta_T^{-1}\right)^{-1} + o_p(1) \\
&\to_d d^{-1}D_2\left(\int_0^1 B_2P'\right)\left(\int_0^1 PP'\right)^{-1}.
\end{aligned}
$$

The last line is obtained directly by (2.8), (2.9), (3.15) and the continuous mapping theorem. The OLS estimator $\widehat{F}$ corresponding to the intercept component is $O_p(T^{1/2})$ and inconsistent. The OLS estimator $\widehat{F}$ corresponding to the deterministic trend component with $s_i \geq 1$ $(i = 2, 3, ..., q)$ is of order $O_p(T^{-s_i+1/2})$. The seasonal differencing of the residual eliminates the intercept component, so

$$
\Delta_d\widehat{u}_{p,t} = D_1\Delta_d x_{1,t} + D_2\Delta_d x_{2,t} + (F - \widehat{F})\Delta_d p_t = D_1\Delta_d x_{1,t} + D_2\Delta_d x_{2,t} + o_p(1),
$$

and hence the bias-correction term in (3.16) can work as usual. If we do not remove the intercept by seasonal differencing, then, in $\widehat{u}_{p,t}$, the deterministic trend component $(F - \widehat{F})p_t$, which is $O_p(T^{1/2})$, dominates both stochastic trend and stationary components and the bias-correction term does not work properly.

To proceed our analysis, we define the weighting matrix by

$$W_T = \begin{bmatrix} W_{1,T} & 0 \\ 0 & W_{2,T} \end{bmatrix} \text{ where } W_{1,T} = T^{1/2} I_{m_1} \text{ and } W_{2,T} = \begin{bmatrix} T I_{m_2} & 0 \\ 0 & T^{1/2}\delta_T \end{bmatrix}.$$

Then,

$$\begin{aligned} \left(\widehat{\Phi}^{+} - \Phi\right) D W_T &= \left[T^{1/2}\left(\widehat{A}_1^{+} - A_1\right) D_1, T\left(\widehat{A}_2^{+} - A_2\right) D_2,\right. \\ &\quad \left. T^{1/2}\left[\left(\widehat{\Pi}^{+} - \Pi\right) + \left(\widehat{A}^{+} - A\right) F\right] \delta_T\right]. \end{aligned}$$

The limiting distribution of $\widehat{\Phi}^{+}$ is derived by the following theorem :

THEOREM 3.5. Under Assumptions 2.1, 2.2, 2.3 and by (3.15),

(a) $\left(\widehat{\Phi}^{+} - \Phi\right) D_L W_{1,T} = T^{1/2}\left(\widehat{A}_1^{+} - A_1\right)$

$$\to_d N\left(0, \left(I \otimes \left(\textstyle\sum_{j=0}^{d-1} \Gamma_{11}(j)\right)^{-1}\right) \Omega_{\varphi\varphi} \left(I \otimes \left(\textstyle\sum_{j=0}^{d-1} \Gamma_{11}'(j)\right)^{-1}\right)\right),$$

(b) $\left(\widehat{\Phi}^{+} - \Phi\right) D_M W_{2,T} = \left[T\left(\widehat{A}_2^{+} - A_2\right) D_2, T^{1/2}\left[\left(\widehat{\Pi}^{+} - \Pi\right) + \left(\widehat{A}^{+} - A\right) F\right] \delta_T\right]$

$$\to_d \left[\int_0^1 dB_{0.2}B_2' \,\vdots\, d\int_0^1 dB_{0.2}p'\right] \begin{bmatrix} d^{-1}\int_0^1 B_2 B_2' & \int_0^1 B_2 p' \\ \int_0^1 p B_2' & \int_0^1 pp' \end{bmatrix}^{-1},$$

where $B_{0.2} = B_0 - \Omega_{02}\Omega_{22}^{-1}B_2 \equiv BM(\Omega_{00.2})$, $\Omega_{00.2} = \Omega_{00} - \Omega_{02}\Omega_{22}^{-1}\Omega_{20}$ and $\Gamma_{11}(j)$ represents the autocovariance of $u_{1,t}$ at the j^{th} period lag. The conditions for both part (a) and (b) to hold are the same as those in Theorem 3.1.

Clearly, we have generalised Theorem 3.1 by including a deterministic trend component. The limiting distribution of the $I(0)$ component is the same as that in Theorem 3.1 (a). The limiting distribution of deterministic trend component is mixed with that of stochastic trend component, but it is still asymptotically mixed normal. This is different from Park and Phillips (1988, Theorem 3.2 and 3.3) where the distribution consists of a mixture of dependent stochastic integrals. By a simple expansion of the result, the limiting distribution of stochastic trend component still depends upon the order of seasonal integration d. Also, the limiting distribution of deterministic component is independent from the parameter d. This is a trivial result of our instrumental variable estimation.

From the previous OLS estimation of F, the intercept component is inconsistent and the limiting distribution of $\widehat{F}$ involves the order of seasonal integration d and unknown transformation matrix D_2. This creates difficulty in drawing inference on the parameters in deterministic trend component. Consequently, we modify the Wald statistic by PFM to draw statistical inference on those parameters in A by a projection of x_t and p_t. Explicitly,

$$\begin{aligned} W_p^+ &= T(Rvec(A)-r)'\left[R\left(I\otimes\left(T^{-1}X'Q_PS(X,d)\right)^{-1}\right)\Omega_{\widehat{\varphi}_x\widehat{\varphi}_x}\right. \\ &\quad \times\left.\left(I\otimes\left(T^{-1}\left(X'Q_PS(X,d)\right)'\right)^{-1}\right)R'\right]^{-1}(Rvec(A)-r), \quad (3.17) \end{aligned}$$

where $Q_P = I - P(P'P)^{-1}P'$. Also, when the condition (3.8) holds,

$$W_p^+ \to_d \chi_q^2.$$

On the other hand, we consider a case where the true model does not contain a deterministic trend component. However, such a component is included in the estimation, that is, $F = 0$. The limiting distribution of $\widehat{\Phi}_2^+$ is not affected as it is free from nuisance parameters in F. The asymptotic variance of $\widehat{A}_2^+$ is still optimal under this overfitting situation. Thus, even if we overfit the deterministic trend component of the model, the asymptotics of $\widehat{\Phi}^+$ are not affected.

4. Asymptotics of FM-SEA Estimators of VAR System

4.1. General Model

In this section, we apply the theorems in previous section to estimate an unrestricted VAR(k) system given by (4.1).

$$y_t = H(L)y_t + v_{0,t}, \qquad for\ t = 1, 2, ..., T, \tag{4.1}$$

where y_t is a n-element column vector, $H(L) = \sum_{j=1}^{k} H_j L^j$ is a $(n \times n)$ matrix polynomial with $k \geq d$. The observations at $t = 0, -1, -2, ..., -d+1$ are initialised by some given distributions.

Now, we write (4.1) in error correction form.

$$y_t = H^*(L)\Delta_d y_t + \sum_{j=1}^{d} \Pi_j m_{j,t-1} + v_{0,t}, \tag{4.2}$$

where the definitions of notations are given by the followings :

When d is odd,

$$\Pi_1 = H(1)/d,$$

$$\Pi_j = \begin{cases} -2\,\mathrm{Re}\left[H[\exp(\iota\theta_{j/2})]\exp(-\iota\theta_{j/2})/\varphi_{j/2}(\exp(\iota\theta_{j/2}))\right] & \text{for } j = 2, 4, ..., d-1, \\ -2\,\mathrm{Re}\left[H[\exp(\iota\theta_{(j-1)/2})]/\varphi_{(j-1)/2}(\exp(\iota\theta_{(j-1)/2}))\right] & \text{for } j = 3, 5, ..., d, \end{cases}$$

$$\varphi_p(L) = \frac{(1-L)\prod_{k=1}^{(d-1)/2}(1 - 2\cos\theta_k L + L^2)}{1 - \exp(-\iota\theta_p)L},$$

$$m_{1,t} = (1 + L + L^2 + ... + L^{d-1})y_t,$$

$$m_{j,t} = \begin{cases} \dfrac{\varphi_{j/2}(L)}{1 - \exp(\iota\theta_{j/2})L} y_t & \text{for } j = 2, 4, ..., d-1, \\ \dfrac{\varphi_{(j-1)/2}(L)L}{1 - \exp(\iota\theta_{(j-1)/2})L} y_t & \text{for } j = 3, 5, ..., d, \end{cases}$$

$$\theta_j = 2j\pi/d;$$

when d is even,

$$\Pi_1 = H(1)/d,\ \Pi_2 = -H(-1)/d,$$

$$\Pi_j = \begin{cases} -2\,\mathrm{Re}\left[H[\exp(\iota\theta_{(j-1)/2})]\exp(-\iota\theta_{(j-1)/2})/\varphi_{(j-1)/2}(\exp(\iota\theta_{(j-1)/2}))\right] \\ \qquad\qquad\qquad\qquad\qquad\qquad \text{for } j = 3, 5, ..., d-1, \\ -2\,\mathrm{Re}\left[H[\exp(\iota\theta_{(j-2)/2})]/\varphi_{(j-2)/2}(\exp(\iota\theta_{(j-2)/2}))\right] \\ \qquad\qquad\qquad\qquad\qquad\qquad \text{for } j = 4, 6, ..., d, \end{cases}$$

$$\varphi_p(L) = \frac{(1-L^2)\prod_{k=1}^{(d-2)/2}(1 - 2\cos\theta_k L + L^2)}{1 - \exp(-\iota\theta_p)L},$$

$$m_{1,t} = (1 + L + L^2 + ... + L^{d-1})y_t,\ m_{2,t} = (1 - L + L^2 - ... - L^{d-1})y_t,$$

$$m_{j,t} = \begin{cases} \dfrac{\varphi_{(j-1)/2}(L)}{1-\exp(\iota\theta_{(j-1)/2})L} y_t & \text{for } j = 3, 5, ..., d-1, \\ \dfrac{\varphi_{(j-2)/2}(L)L}{1-\exp(\iota\theta_{(j-2)/2})L} y_t & \text{for } j = 4, 6, ..., d, \end{cases}$$

$\theta_j = 2j\pi/d.$

Also, (4.2) can written as

when d is odd,

$$\begin{aligned} \Delta_d y_t \;\; = \;\; & H^*(L)\Delta_d y_t + (\Pi_1 - I/d)\, m_{1,t-1} \\ & + \sum_{j=1}^{(d-1)/2} (2\,\mathrm{Re}\,[\exp(-\iota\theta_j)/\varphi_j(\exp(-\iota\theta_j))]\, I + \Pi_{2j})\, m_{2j,t-1} \\ & + \sum_{j=1}^{(d-1)/2} (2/\,\mathrm{Re}\,[\varphi_j(\exp(-\iota\theta_j))]\, I + \Pi_{2j+1})\, m_{2j+1,t-1} + v_{0,t}; \end{aligned} \quad (4.3)$$

when d is even,

$$\begin{aligned} \Delta_d y_t \;\; = \;\; & H^*(L)\Delta_d y_t + (\Pi_1 - I/d)\, m_{1,t-1} + (\Pi_2 + I/d)\, m_{2,t-1} \\ & + \sum_{j=1}^{(d-2)/2} (2\,\mathrm{Re}\,[\exp(-\iota\theta_j)/\varphi_j(\exp(-\iota\theta_j))]\, I + \Pi_{2j+1})\, m_{2j+1,t-1} \\ & + \sum_{j=1}^{(d-2)/2} (2/\,\mathrm{Re}\,[\varphi_j(\exp(-\iota\theta_j))]\, I + \Pi_{2j})\, m_{2j,t-1} + v_{0,t}. \end{aligned} \quad (4.4)$$

The error correction forms (4.3) and (4.4) are more general than those suggested by HEGY (p.232) and Lee (1992, p.7). Equations (4.3) and (4.4) allow for more general ECM terms, which describe the long-run equilibrium relationships of elements of y_t, at different frequencies. Although the expressions at the first glance are quite complicated, a simpler form will be presented in the next section. Note that the variables $m_{1,t}$, $m_{2,t}$, ..., $m_{d,t}$ in (4.2) and the products are filtered values of y_t such that they have unit roots at frequency $0, j/d$ for $j = 1, 2, ..., (d-1)/2$ respectively when d is odd or at frequency $0, 1/2$ and j/d for $j = 1, 2, ..., (d-2)/2$ respectively when d is even. Also, the coefficient matrices $\Pi_1, \Pi_2, ..., \Pi_d$ contain the information of long-run equilibrium relationships of elements in y_t. This implies that, in (4.3), the products $(\Pi_1 - I/d)\, m_{1,t-1}$, $(2\,\mathrm{Re}\,[\exp(-\iota\theta_j)/\varphi_j(\exp(-\iota\theta_j))]\, I + \Pi_{2j})\, m_{2j,t-1}$ and $(2/\,\mathrm{Re}\,[\varphi_j(\exp(-\iota\theta_j))]\, I + \Pi_{2j+1})\, m_{2j+1,t-1}$ and, in (4.4), the products $(\Pi_1 - I/d)\, m_{1,t-1}$, $(\Pi_2 + I/d)\, m_{2,t-1}$, $(2\,\mathrm{Re}\,[\exp(-\iota\theta_j)/\varphi_j(\exp(-\iota\theta_j))]\, I + \Pi_{2j+1})$ $\times m_{2j+1,t-1}$ and $(2/\,\mathrm{Re}\,[\varphi_j(\exp(-\iota\theta_j))]\, I + \Pi_{2j})\, m_{2j,t-1}$ are the error correction mechanism. The terms $\Delta_d y_{t-j}$ for $j = 1, 2, ..., k-d$ indicate the short-run dynamics. Combining all together, (4.3) and (4.4) describe the adjustment of the systems in the presence of deviation of long-run equilibrium. To proceed our analysis, we make the following assumptions :

ASSUMPTION 4.1.

(a) When d is odd, $\Pi_1 = \alpha_1\beta_1' + I/d$,

$\Pi_j = \alpha_j\beta_{j/2+1}' - 2\,\mathrm{Re}\,[\exp(-\iota\theta_j)/\varphi_j(\exp(-\iota\theta_j))]\, I$ for $j = 2, 4, ..., d-1$, $\Pi_j = \alpha_j\beta_{(j+1)/2}' - 2/\,\mathrm{Re}\,[\varphi_j(\exp(-\iota\theta_j))]\, I$ for $j = 3, 5, ..., d$ where α_1 is of order $(n \times r_1)$, α_j and α_{j+1} are of order $(n \times r_j)$ for $j = 2, 4, ..., d-1$. Also, β_j is of order $(n \times r_j)$ for $j = 1, 2, ..., (d+1)/2$. The matrices α_j's and β_j's have the full column rank, $0 \leq r_j \leq n$ for $j \geq 1$.

(b) When d is even, $\Pi_1 = \alpha_1\beta_1' + I/d$, $\Pi_2 = \alpha_2\beta_2' - I/d$, $\Pi_j = \alpha_j\beta_{(j+3)/2}' - 2\,\mathrm{Re}\left[\exp(-\iota\theta_j)/\varphi_j(\exp(-\iota\theta_j))\right] I$ for $j = 3, 5, ..., d-1$, $\Pi_j = \alpha_j\beta_{(j+2)/2}' - 2/\,\mathrm{Re}\left[\varphi_j(\exp(-\iota\theta_j))\right] I$ for j $= 4, 6, ..., d$ where α_1 is of order $(n \times r_1)$, α_2 is of order $(n \times r_2)$, α_j and α_{j+1} are of order $(n \times r_j)$ for $j = 3, 5, ..., d-1$. Also, β_j is of order $(n \times r_j)$ for $j = 1, 2, ..., (d+2)/2$. α_j's and β_j's have the full column rank, $0 \leq r_j \leq n$ for $j \geq 1$.

(c) When d is odd, $\alpha_{1\perp}$ is a matrix of order $(n \times (n-r_1))$, $\alpha_{j\perp}$ and $\alpha_{j+1\perp}$ are of order $(n \times (n-r_j))$ for $j = 2, 4, ..., d-1$. Also, $\beta_{j\perp}$ is of order $(n \times (n-r_j))$ for $j = 1, 2, ..., (d+1)/2$. They should satisfy the conditions $\alpha_{j\perp}'\alpha_j = 0$ for $j = 1, 2, ..., d$, $\beta_{j\perp}'\beta_j = 0$ for $j = 1, 2, ..., (d+1)/2$. When $r_j = 0$ for $j = 1, 2, ..., (d+1)/2$, we take $\beta_{1\perp} = \beta_{2\perp} = ... = \beta_{(d+1)/2\perp} = I$.

(d) When d is even, $\alpha_{1\perp}$ and $\alpha_{2\perp}$ are matrices of order $(n \times (n-r_1))$ and $(n \times (n-r_2))$ respectively, $\alpha_{j\perp}$ and $\alpha_{j+1\perp}$ are of order $(n \times (n-r_j))$ for $j = 3, 5, ..., d-1$. Also, $\beta_{j\perp}$ is of order $(n \times (n-r_j))$ for $j = 1, 2, ..., (d+2)/2$. They should satisfy the conditions $\alpha_{j\perp}'\alpha_j = 0$ for $j = 1, 2, ..., d$, $\beta_{j\perp}'\beta_j = 0$ for $j = 1, 2, ..., (d+2)/2$. When $r_j = 0$ for $j = 1, 2, ..., (d+2)/2$, we take $\beta_{1\perp} = \beta_{2\perp} = ... = \beta_{(d+2)/2\perp} = I$.

Since a conjugate pair of complex unit roots shares the same frequency, we cannot distinguish the cointegrating vectors corresponding to the complex unit roots in a conjugate pair separately. Thus, we assume that they have the same cointegrating matrix β_j for all $j \geq 2$ when d is odd and for all $j \geq 3$ when d is even. We rewrite (4.2) in a more compact form

$$y_t = H^* z_t + \Pi p_{t-1} + v_{0,t} = A x_t + v_{0,t}, \tag{4.5}$$

where $H^* = \left[H_1^*, H_2^*, ..., H_{k-d}^*\right]$, $z_t = \left(\Delta_d y_{t-1}', \Delta_d y_{t-2}', ..., \Delta_d y_{t-k+d}'\right)'$, $\Pi = [\Pi_1, \Pi_2, ..., \Pi_d]$, $p_{t-1} = \left(m_{1,t-1}', m_{2,t-1}', ..., m_{d,t-1}'\right)'$, $A = [H^*, \Pi]$ and $x_t = \left(z_t', p_{t-1}'\right)'$. Next, we define the transformation matrix D for z_t and p_{t-1} by

$$D = \begin{bmatrix} D_L, & D_M, & D_N \end{bmatrix}, D_L = \begin{bmatrix} I_{k-d}, & 0 \end{bmatrix}',$$

$$D_M = \begin{cases} diag\left(\beta_1, \beta_2, \beta_2, \beta_3, \beta_3, ..., \beta_{(d+1)/2}, \beta_{(d+1)/2}\right) & \text{for } d \text{ is odd,} \\ diag\left(\beta_1, \beta_2, \beta_3, \beta_3, ..., \beta_{(d+2)/2}, \beta_{(d+2)/2}\right) & \text{for } d \text{ is even,} \end{cases}$$

$$D_N = \begin{cases} diag\left(\beta_{1\perp}, \beta_{2\perp}, \beta_{2\perp}, \beta_{3\perp}, \beta_{3\perp}, ..., \beta_{(d+1)/2\perp}, \beta_{(d+1)/2\perp}\right) \\ \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \text{for } d \text{ is odd,} \\ diag\left(\beta_{1\perp}, \beta_{2\perp}, \beta_{3\perp}, \beta_{3\perp}, ..., \beta_{(d+2)/2\perp}, \beta_{(d+2)/2\perp}\right) \\ \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \text{for } d \text{ is even.} \end{cases}$$

Now, we decompose p_{t-1} in (4.5) by premultiplying $[D_M, D_N]\,[D_M, D_N]'$ on p_{t-1}.

$$y_t = H^* z_t + \Pi_M p_{1,t-1} + \Pi_N p_{2,t-1} + v_{0,t}, \tag{4.6}$$

where $\Pi_M = \Pi D_M$, $\Pi_N = \Pi D_N$, $D_M' p_{t-1} = p_{1,t-1} = v_{1,t}$ and $D_N' \Delta_d p_{t-1} = \Delta_d p_{2,t-1} = v_{2,t}$. Clearly, when d is odd,

$$\Pi_N = diag\left(\beta_{1\perp}/d, \beta_{2\perp}\Pi_{N,2}, \beta_{3\perp}\Pi_{N,3}, ..., \beta_{d\perp}\Pi_{N,d}\right), \tag{4.7}$$

$$\text{where } \Pi_{N,j} = \begin{cases} -2\operatorname{Re}\left[\exp(-\iota\theta_j)/\varphi_j(\exp(-\iota\theta_j))\right] I & \text{for } j = 2, 4, ..., d-1, \\ -2/\operatorname{Re}\left[\varphi_j(\exp(-\iota\theta_j))\right] I & \text{for } j = 3, 5, ..., d. \end{cases}$$

When d is even,

$$\Pi_N = diag\left(\beta_{1\perp}/d, -\beta_{2\perp}/d, \beta_{3\perp}\Pi_{N,3}, \beta_{4\perp}\Pi_{N,4}, ..., \beta_{d\perp}\Pi_{N,d}\right), \tag{4.8}$$

$$\text{where } \Pi_{N,j} = \begin{cases} -2\operatorname{Re}\left[\exp(-\iota\theta_j)/\varphi_j(\exp(-\iota\theta_j))\right] I & \text{for } j = 3, 5, ..., d-1, \\ -2/\operatorname{Re}\left[\varphi_j(\exp(-\iota\theta_j))\right] I & \text{for } j = 4, 6, ..., d, \end{cases}$$

by Assumption 4.1 (b). Next, we define $v_t = \left(v_{0,t}', v_{1,t}', v_{2,t}'\right)'$ and $\varphi_t = v_{0,t} \otimes v_{1,t}$ and they satisfy the following assumption :

ASSUMPTION 4.2.

(a) $v_t = C(L)\varepsilon_t = \sum_{j=0}^{\infty} C_j \varepsilon_{t-j}$, $\sum_{j=0}^{\infty} j^a \|C_j\| < \infty$ for some constant $a > 1$, $|C(1)| \neq 0$;

(b) ε_t is i.i.d. with zero mean, variance matrix $\Sigma_{\varepsilon\varepsilon} > 0$ and finite fourth order cumulants;

(c) Let $\xi_t = \left(v_{0,t}', v_{1,t+1}'\right)'$ and $\Im_t = \sigma\left(\xi_t, \xi_{t-1}, ..., \xi_1\right)$ be the σ-field generated by $\{\xi_j\}_1^t$. Then $(v_{0,t}, \Im_t)$ is a martingale difference sequence, that is, $E(v_{0,t} \mid \Im_{t-1}) = 0$.

Assumption 4.2 (c) ensures that $E(v_{0,t+j}v_{1,t}') = 0$ for all $j \geq 0$ by iterative expectation. This also implies that $E(\varphi_{t,j}) = E(v_{0,t+j} \otimes v_{1,t}) = 0$ for all $j \geq 0$ because the assumption of martingale difference sequence for $v_{0,t}$ and $v_{1,t}$ is stronger than the assumption of no serial correlation as in Assumption 2.1 (c). Therefore, under Assumption 4.2 (c), the autocovariance of φ_t at the j^{th} period lag is given by

$$E(\varphi_t \varphi_{t+j}') = E(v_{0,t}v_{0,t+j}' \otimes v_{1,t}v_{1,t+j}') = \begin{cases} 0 & \text{for all } j \neq 0, \\ \Sigma_{00} \otimes \Sigma_{11} & \text{for } j = 0, \end{cases} \tag{4.9}$$

where Σ_{00} and Σ_{11} are the variance matrices of $v_{0,t}$ and $v_{1,t}$ respectively. Furthermore, we define and partition the long-run covariance and one-sided long-run covariance matrices of v_t by

$$\Omega = \begin{bmatrix} \Omega_{00} & \Omega_{01} & \Omega_{02} \\ \Omega_{10} & \Omega_{11} & \Omega_{12} \\ \Omega_{20} & \Omega_{21} & \Omega_{22} \end{bmatrix} \text{ and } \Delta = \begin{bmatrix} \Delta_{00} & \Delta_{01} & \Delta_{02} \\ \Delta_{10} & \Delta_{11} & \Delta_{12} \\ \Delta_{20} & \Delta_{21} & \Delta_{22} \end{bmatrix}.$$

The definitions of Ω and Δ are analogous to those in Section 2. From (4.6), we define $x_{1,t} = \left(z_t', p_{1,t-1}'\right)'$ and $x_{2,t} = p_{2,t-1}$. Then, (4.6) is rewritten as

$$y_t = A_1 x_{1,t} + A_2 x_{2,t} + v_{0,t}, \tag{4.10}$$

where $A_1 = [H^*, \Pi_M]$ and $A_2 = \Pi_N$. Now, we express (4.10) in data matrix form.

$$Y' = A_1 X_1' + A_2 X_2' + V_0'. \tag{4.11}$$

By (3.3), the FM-SEA regression on (4.5) yields the estimates by the following formula :

$$\widehat{A}^+ = \left[\widehat{H}^{*+}, \widehat{\Pi}^+\right] = \left(Y^{+\prime} S(X,d) - T\widehat{\Delta}_{0x}^+\right) \left(X' S(X,d)\right)^{-1}, \tag{4.12}$$

where $\widehat{A}_1^+ = \widehat{A}^+ \left[D_L, D_M\right]$, $\widehat{A}_2^+ = \widehat{A}^+ D_N$, $X' = \left(Z', P_{-1}'\right)'$, Z and P_{-1} are the data matrices of z_t and p_{t-1} respectively, $\widehat{\Delta}_{0x}^+ = \widehat{\Delta}_{0x} - \widehat{\Omega}_{0x}\widehat{\Omega}_{xx}^{-1}\widehat{\Delta}_{xx}$, Δ_{0x} and Ω_{0x} are the one-sided long-run covariance and long-run covariance matrices of $(\widehat{v}_{0,t}, \Delta_d x_t)$ respectively, Δ_{pp} and Ω_{pp} are the one-sided long-run covariance and long-run covariance matrices of $(\Delta_d x_t, \Delta_d x_t)$ respectively. Furthermore, the kernel estimates $\widehat{\Omega}_{0x}$, $\widehat{\Omega}_{xx}$, $\widehat{\Delta}_{0x}$ and $\widehat{\Delta}_{xx}$ are obtained by OLS estimation of (4.5).

We observe that the endogeneity correction term is still present in the estimation of $\widehat{H}^{*+}$ in (4.12) although we realise that z_t is an $I(0)$ variable. The inclusion of this endogeneity correction term for the estimation of $\widehat{H}^{*+}$ is important to the asymptotics of $\widehat{A}_1^+$. Also, the transformation of variable z_t ensures the consistency of $\widehat{H}^{*+}$.

The limiting distributions of $\widehat{A}_1^+$ and $\widehat{A}_2^+$ in (4.12) are given by the following theorem :

THEOREM 4.1. Under Assumptions 2.2, 2.3, 4.1 and 4.2,

(a) $T^{1/2}\left(\widehat{A}_1^+ - A_1\right) = T^{1/2}\left[\left(\widehat{H}^{*+} - H^*\right), \left(\widehat{\Pi}_M^+ - \Pi_M\right)\right]$

$$\rightarrow_d N\left(0, \Sigma_{00} \otimes \left(\textstyle\sum_{j=0}^{d-1}\Gamma_{11}(j)\right)^{-1}\Sigma_{11}\left(\textstyle\sum_{j=0}^{d-1}\Gamma_{11}'(j)\right)^{-1}\right),$$

(b) $T\left(\widehat{A}_2^+ - A_2\right) = T\left(\widehat{\Pi}_N^+ - \Pi_N\right) \rightarrow_d d\left(\int_0^1 dB_{0.2}B_2'\right)\left(\int_0^1 B_2 B_2'\right)^{-1}$,

where $\Sigma_{00} = E(v_{0,t}v_{0,t}')$, $\Sigma_{11} = E(x_{1,t}x_{1,t}')$, $\Gamma_{11}(j) = E(x_{1,t}x_{1,t-j}')$, $B_{0.2} = B_0 - \Omega_{02}\Omega_{22}^{-1}B_2 \equiv BM(\Omega_{00.2})$, $\Omega_{00.2} = \Omega_{00} - \Omega_{02}\Omega_{22}^{-1}\Omega_{20}$. Part (a) holds when Assumption 2.3 (b) is satisfied while part (b) holds when Assumption 2.3 (c) is satisfied. For both part (a) and (b) to hold, Assumption 2.3 (a) should be satisfied.

This theorem tells us that FM-SEA regression on an unrestricted VAR system does not require the information on the number of cointegrating vectors. In addition, the distribution of FM-SEA estimator is asymptotically normal irrespective to the cointegrating rank while the limiting distribution of OLS estimator in cointegrated system, in Phillips and Durlauf (1986, Theorem 3.2), is the unit root type. Furthermore, unlike Engle and Granger (1987)'s two-step procedure, the FM-SEA estimation of (4.2), as in Phillips (1991, equation (7)), does not need to estimate the cointegrating matrix. On the other hand, as mentioned before, the limiting distribution of non-cointegrated component changes with the order of seasonal integration. When $d = 1$, the asymptotic variance of $\widehat{A}_1^+$ is reduced to the result of PFM, $\Sigma_{00} \otimes \Sigma_{11}^{-1}$.

COROLLARY 4.2. When no element of y_t is seasonally cointegrated, that is, the number of columns of D_M is zero, then under Assumptions 2.2, 2.3 (d), 4.1 and 4.2, the followings hold :

(a) $T^{1/2}\left(\widehat{A}_1^+ - A_1\right) = T^{1/2}\left(\widehat{H}^{*+} - H^*\right)$

$$\rightarrow_d N\left(0, \Sigma_{00} \otimes \left(\textstyle\sum_{j=0}^{d-1} \Gamma_{11}(j)\right)^{-1} \Sigma_{11} \left(\textstyle\sum_{j=0}^{d-1} \Gamma_{11}'(j)\right)^{-1}\right),$$

(b) $T\left(\widehat{A}_2^+ - A_2\right) = T\left(\widehat{\Pi}^+ - \Pi\right) \rightarrow_p 0,$

where $\Pi = \Pi_N$ has been defined by (4.7) and (4.8). In this corollary, $x_{1,t}$ contains the lagged values of $\Delta_d y_t$ only. $x_{2,t}$ contains the non-stationary component $p_{2,t-1}$ with elements unit roots at various frequencies. Hence, we can estimate the model with seasonally differenced variables as regressors by OLS. Nevertheless, FM-SEA estimation provides more information about the non-stationary component. Note that Π is different from that in PFM because the seasonally integrated series contain the unit roots at various frequencies while the usual $I(1)$ series contain the unit root at frequency zero only.

COROLLARY 4.3. When all elements of y_t are seasonally cointegrated, that is, the number of columns of D_N is zero, then under Assumptions 2.2, 2.3 (b), 4.1 and 4.2, the following holds :

$$T^{1/2}\left(\widehat{A}_1^+ - A_1\right) \rightarrow_d N\left(0, \Sigma_{00} \otimes \left(\sum_{j=0}^{d-1} \Gamma_{11}(j)\right)^{-1} \Sigma_{11} \left(\sum_{j=0}^{d-1} \Gamma_{11}'(j)\right)^{-1}\right).$$

In this corollary, the regressor x_t contains only stationary component z_t and the cointegrated component $p_{1,t-1}$, which equals p_{t-1} in this case. In this cointegrated system, FM-SEA estimation can obtain a more general result than Phillips (1991).

Suppose we test the hypothesis $H_0 : Rvec(A) = r$ where the definitions of R and r are the same as those in Section 3.1. Then, the Wald statistic can be followed easily from (3.7).

$$\begin{aligned} W_V^+ &= T(Rvec(A)-r)' \left[R \left(\hat{\Sigma}_{00} \otimes \left(T^{-1}X'S(X,d)\right)^{-1} \left(T^{-1}X'X\right) \right.\right. \\ &\quad \left.\left. \times \left(T^{-1}\left(X'S(X,d)\right)'\right)^{-1} \right) R' \right]^{-1} (Rvec(A)-r). \end{aligned} \tag{4.13}$$

If the condition

$$rank \left[R \left\{ \Sigma_{00} \otimes \left(\sum_{j=0}^{d-1} \Gamma_{11}(j) \right)^{-1} \Sigma_{11} \left(\sum_{j=0}^{d-1} \Gamma'_{11}(j) \right)^{-1} \right\} R' \right] = q, \tag{4.14}$$

holds, then

$$W_V^+ \to_d \chi_q^2.$$

4.2. Model with d = 4

In this section, we study the model with $d = 4$ because many economic data are observed quarterly. Now, we write (4.1) in error correction form.

$$y_t = H^*(L)\Delta_4 y_t + \Pi_1 m_{1,t-1} + \Pi_2 m_{2,t-1} + \Pi_3 m_{3,t-1} + \Pi_4 m_{4,t-1} + v_{0,t}, \quad (4.15)$$

or

$$\begin{aligned}\Delta_4 y_t &= H^*(L)\Delta_4 y_t + (\Pi_1 - I/4)\, m_{1,t-1} + (\Pi_2 + I/4)\, m_{2,t-1} + \Pi_3 m_{3,t-1} \\ &\quad + (\Pi_4 + I/2)\, m_{4,t-1} + v_{0,t}, \qquad (4.16)\end{aligned}$$

where $m_{1,t} = (1 + L + L^2 + L^3)y_t$, $m_{2,t} = (1 - L + L^2 - L^3)y_t$, $m_{3,t} = (1 - L^2)y_t$, $m_{4,t} = L(1 - L^2)y_t$, $\Pi_1 = H(1)/4$, $\Pi_2 = -H(-1)/4$, $\Pi_3 = -\operatorname{Re}[-H(\iota)\iota]/2 = -\operatorname{Im}[H(\iota)]/2$, $\Pi_4 = -\operatorname{Re}[H(\iota)]/2$, $H^*(L) = \sum_{i=1}^{k-4} H_i^* L^i$ and $H_i^* = -\sum_{j=1}^{[(k-i)\,/\,4]} H_{i+4j}$.

This error correction form is the one suggested by HEGY (p.232) and Lee (1992, p.7). Obviously, the variables $m_{1,t}$, $m_{2,t}$ and $m_{3,t}$ are filtered values of y_t that have unit roots at frequency $0, 1/2$ and $1/4$ respectively. Furthermore, the transformation matrix is defined by $D = [D_L, D_M, D_N]$, $D_L = [I_{k-4}, 0]'$, $D_M = diag\,(\beta_1, \beta_2, \beta_3, \beta_3)$ and $D_N = diag\,(\beta_{1\perp}, \beta_{2\perp}, \beta_{3\perp}, \beta_{3\perp})$. Also, from (4.8), $\Pi_N = diag\,(\beta_{1\perp}/4, -\beta_{2\perp}/4, 0, -\beta_{3\perp}/2)$.

Then, we can estimate (4.15) by using (4.12)

$$\widehat{A}^+ = \left[\widehat{H}^{*+}, \widehat{\Pi}^+\right] = \left(Y^{+\prime} S(X,4) - T\widehat{\Delta}_{0x}^+\right)\left(X' S(X,4)\right)^{-1}, \quad (4.17)$$

where $\widehat{\Delta}^{+}_{0x} = \widehat{\Delta}_{0x} - \widehat{\Omega}_{0x}\widehat{\Omega}^{-1}_{xx}\widehat{\Delta}_{xx}$, Δ_{0x}, Ω_{0x}, Δ_{xx} and Ω_{xx} are defined in Section 4.1 with $d = 4$. Furthermore, the kernel estimates $\widehat{\Omega}_{0x}$, $\widehat{\Omega}_{xx}$, $\widehat{\Delta}_{0x}$ and $\widehat{\Delta}_{xx}$ can be obtained by OLS estimation of (4.15).

The limiting distribution of $\widehat{A}^{+}$ in (4.17) is given by the following results :

COROLLARY 4.4. Under the conditions of Theorem 4.1 with $d = 4$, then

(a) $T^{1/2}\left(\widehat{A}^{+}_{1} - A_{1}\right) \rightarrow_{d} N\left(0, \Sigma_{00} \otimes \left(\sum_{j=0}^{3}\Gamma_{11}(j)\right)^{-1}\Sigma_{11}\left(\sum_{j=0}^{3}\Gamma'_{11}(j)\right)^{-1}\right)$,

(b) $T\left(\widehat{A}^{+}_{2} - A_{2}\right) = T\left(\widehat{\Pi}^{+}_{N} - \Pi_{N}\right) \rightarrow_{d} 4\left(\int_{0}^{1} dB_{0.2}B'_{2}\right)\left(\int_{0}^{1} B_{2}B'_{2}\right)^{-1}$.

The conditions for this corollary to hold are the same as those in Theorem 4.1.

Obviously, the result of this corollary is a special case of Theorem 4.1. One interesting finding is obtained in the non-stationary component. If all variables in y_t are cointegrated, then, by Corollary 4.2, $\widehat{\Pi}^{+}_{N}$ converges to Π_N. For the cointegrated component at frequency zero, the corresponding element of Π_N is $I/4$ by (4.8). This result is quite interesting because, in a usual I(1) cointegrated system, the corresponding element of Π_N is I. This can be explained intuitively by dividing the unit circle into four equal portions. Then, frequency zero and 1/2 components are represented by the upper right and upper left portions respectively. The frequency 1/4 component is allocated to the lower half of the unit circle. Of course, for the I(1) process, it will get all the unit circle at one time.

EXAMPLE 4.1. To illustrate the use of theorem derived, we consider a special case of model (4.1) with $n = 1$ and $k = 4$. In particular,

$$y_t = \sum_{j=1}^{4} H_j y_{t-j} + v_{0,t}. \tag{4.18}$$

We rewrite (4.18) as (4.5), then

$$y_t = \Pi p_{t-1} + v_{0,t}. \tag{4.19}$$

Applying the formula (4.18),

$$\widehat{A}^{+} = \widehat{\Pi}^{+} = \left[Y^{+\prime}S(P_{-1}, 4) - T\widehat{\Delta}_{0p}^{+}\right]\left(P_{-1}^{\prime}S(P_{-1}, 4)\right)^{-1}. \tag{4.20}$$

where

$$\begin{aligned}Y^{+\prime}S(P_{-1}, 4) = \Big[&\textstyle\sum_{t=2}^{T} y_t^{+} S(m_{1,t-1}, 4), \sum_{t=2}^{T} y_t^{+} S(m_{2,t-1}, 4), \sum_{t=2}^{T} y_t^{+} S(m_{3,t-1}, 4),\\ &\textstyle\sum_{t=2}^{T} y_t^{+} S(m_{4,t-1}, 4)\Big]\end{aligned}$$

and

$$P_{-1}^{\prime}S(P_{-1}, 4) = \begin{bmatrix} \sum_{t=2}^{T} m_{1,t-1}S(m_{1,t-1}, 4) & \sum_{t=2}^{T} m_{1,t-1}S(m_{2,t-1}, 4) & \sum_{t=2}^{T} m_{1,t-1}S(m_{3,t-1}, 4) & \sum_{t=2}^{T} m_{1,t-1}S(m_{4,t-1}, 4) \\ \sum_{t=2}^{T} m_{2,t-1}S(m_{1,t-1}, 4) & \sum_{t=2}^{T} m_{2,t-1}S(m_{2,t-1}, 4) & \sum_{t=2}^{T} m_{2,t-1}S(m_{3,t-1}, 4) & \sum_{t=2}^{T} m_{2,t-1}S(m_{4,t-1}, 4) \\ \sum_{t=2}^{T} m_{3,t-1}S(m_{1,t-1}, 4) & \sum_{t=2}^{T} m_{3,t-1}S(m_{2,t-1}, 4) & \sum_{t=2}^{T} m_{3,t-1}S(m_{3,t-1}, 4) & \sum_{t=2}^{T} m_{3,t-1}S(m_{4,t-1}, 4) \\ \sum_{t=2}^{T} m_{4,t-1}S(m_{1,t-1}, 4) & \sum_{t=2}^{T} m_{4,t-1}S(m_{2,t-1}, 4) & \sum_{t=2}^{T} m_{4,t-1}S(m_{3,t-1}, 4) & \sum_{t=2}^{T} m_{4,t-1}S(m_{4,t-1}, 4) \end{bmatrix}$$

Then, we have the following corollary :

COROLLARY 4.5. Under the conditions of Theorem 4.1 with $d = 4$, $n = 1$ and $k = 4$, we have

$$\text{(a) } T^{1/2}\left(\widehat{A}_1^{+} - A_1\right) = T^{1/2}\left(\widehat{\Pi}_M^{+} - \Pi_M\right) \to_d N\left(0, \Sigma_{00}\Sigma_{11}/\left(\Sigma_{j=0}^{3}\right)^2\right),$$

$$\text{(b) } T\left(\widehat{A}_2^{+} - A_2\right) = T\left(\widehat{\Pi}_N^{+} - \Pi_N\right) \to_d 4\left(\int_0^1 dB_{0.2}B_2'\right)\left(\int_0^1 B_2B_2'\right)^{-1},$$

where the variance matrices Σ_{00}, Σ_{11} and autocovariances $\Gamma_{11}(j)$ are scalar in this single equation model. The conditions for this corollary to hold are the same as those in Theorem 4.1.

In this corollary, the ECM model contains the cointegrated and non-cointegrated components, no short-run dynamics. The result of part (a) is different from that by OLS estimation of single equation due to the transformation $S(., d)$ on the regressors. Also, in this AR(4) model, we also obtain the limiting distribution of non-cointegrated component because $SI(4, \theta)$ series are cointegrated at three different frequencies. Despite the presence of non-stationary component, we do not need to perform the unit root tests (See HEGY and Dickey et al (1984)) because of the dominance of cointegrated component.

5. Monte Carlo Experimental Results

In previous sections, we derive the results asymptotically. This section studies the behaviour of FM-SEA estimator in finite samples. Phillips and Hansen (1990) compared the finite sample properties of FM-OLS, OLS and ECM estimators for a model with an $I(1)$ process. In general, the FM-OLS and ECM estimators perform better than OLS estimators. Although there is a persistent and smaller bias in the FM-OLS estimator, FM-OLS estimator is superior to the ECM estimator. Kitamura and Phillips (1995) compared the finite sample properties of FM-IV, FM-GIVE, FM-GMM, FM-OLS and OLS estimators for a model with a mixture of $I(0)$ and $I(1)$ processes. The basic result is that the FM-IV estimator has the best overall performance among those chosen estimators, although it has a greater root mean squared error. In the followings, we will study the finite sample properties of FM-SEA and OLS estimators with a mixture of $I(0)$ and $SI(d,\theta)$ processes.

Now, let us construct a single equation model.

$$y_t = a'x_t + u_{0,t}, \tag{5.1}$$

where x_t and $u_{0,t}$ are generated by

$$x_t = \begin{bmatrix} x_{1,t} \\ x_{2,t} \end{bmatrix} = \begin{bmatrix} \lambda & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x_{1,t-d} \\ x_{2,t-d} \end{bmatrix} + \begin{bmatrix} u_{a,t} \\ u_{b,t} \end{bmatrix}, \ |\lambda| < 1, \tag{5.2}$$

$$u_{0,t} = \mu u_{b,t} + u_{c,t}, \tag{5.3}$$

$$\begin{bmatrix} u_{a,t} \\ u_{b,t} \\ u_{c,t} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \varepsilon_{a,t} \\ \varepsilon_{b,t} \\ \varepsilon_{c,t} \end{bmatrix} + \begin{bmatrix} \rho & 0 & 0 \\ 0 & \rho & 0 \\ 0 & 0 & \rho \end{bmatrix} \begin{bmatrix} \varepsilon_{a,t-1} \\ \varepsilon_{b,t-1} \\ \varepsilon_{c,t-1} \end{bmatrix}. \tag{5.4}$$

From (5.2), we can see that $x_{1,t}$ and $x_{2,t}$ are generated by restricted AR(d) and $SI(d,\theta)$ processes respectively. For the generation process of regression error $u_{0,t}$, μ indicates the degree of contemporaneous correlation between $x_{2,t}$ and $u_{0,t}$. In (5.4), the components $u_{a,t}$, $u_{b,t}$ and $u_{c,t}$ are generated by independent MA(1) processes because $\varepsilon_{a,t}$, $\varepsilon_{b,t}$ and $\varepsilon_{c,t}$ are i.i.d. in our model. We also assume that they are generated by the same parameter ρ for simplicity. Furthermore, the error component $u_{a,t}$ is serially uncorrelated with the regression error $u_{0,t}$. This makes the error components in this simulation model to satisfy Assumption 2.1.

The values of $\varepsilon_{a,t}$, $\varepsilon_{b,t}$, $\varepsilon_{c,t}$ and x_t at $t = 0, 1, 2, ..., d-1$ are excluded in our simulation model and we choose $a' = (a_1, a_2) = (5, 10)$, $\lambda = 0.9$ and $\rho = 0.9$. The results are based on 20,000 simulations of sample size $T = 50, 100, 200$ and 500. Furthermore, $d = 1$ and 4 are chosen for comparison of the results. The FM-SEA estimators of a_1 and a_2 are given by (3.10) and (3.11) respectively. The OLS estimators of a_1 and a_2 are given by

$$\begin{bmatrix} \widehat{a}_1 \\ \\ \widehat{a}_2 \end{bmatrix} = \begin{bmatrix} \sum_{t=1}^{T} x_{1,t}^2 & \sum_{t=1}^{T} x_{1,t}x_{2,t} \\ \\ \sum_{t=1}^{T} x_{1,t}x_{2,t} & \sum_{t=1}^{T} x_{2,t}^2 \end{bmatrix}^{-1} \begin{bmatrix} \sum_{t=1}^{T} x_{1,t}y_t \\ \\ \sum_{t=1}^{T} x_{2,t}y_t \end{bmatrix}. \tag{5.5}$$

Parzen kernel is chosen arbitrarily for the estimation of long-run covariance matrices. Kitamura and Phillips (1995) mentioned that there was no reason to choose optimal value of K for the estimation of a_1 and a_2. We choose the lag length K of kernel estimation of Δ and Ω to be 8 arbitrarily throughout this simulation. (The results change very small for the other values of K, so we will not report the results here). The simulation results are summarised in the diagrams and tables which are given at the end of the appendix. Before we proceed to analyse the simulation results, we define some notations for convenience. MAD, Bias_{avg} and RMSE_{avg} denote the mean absolute deviation, average bias and average root mean squared error respectively. The probability density functions (pdf's) of $\widehat{a}_1^+$ and $\widehat{a}_2^+$ are presented selectively and all results of summary statistics are shown in Tables 1 - 4.

Figures 1 and 2 show the pdf's of estimates of coefficients a_1 and a_2 respectively. In those figures, we use the bandwidths of 0.03 and 0.01 for pdf's under $T = 50$ and $T = 500$ respectively. We observe that the pdf's of $\widehat{a}_1^+$ are flatter than those of $\widehat{a}_1$ under various values of μ, d and T in Figure 1. This can be explained by two reasons. First, we realise that $\Delta_d x_{1,t}$ is a moving average process.

This induces the loss of information of $x_{1,t}$ in finite samples because the long-run variance matrix for this over-differenced series $\Delta_d x_{1,t}$ is zero. Second, the kernel estimates of long-run covariance matrices do not perform well in finite samples. When the sample size T increases, the peaks of pdf's of $\widehat{a}_1^+$ and $\widehat{a}_1$ become closer. This represents that the FM-SEA and OLS estimators converge together asymptotically. As d increases, the pdf's get flatter because the variances of $\widehat{a}_1^+$ and $\widehat{a}_1$ are related to the values of d by Corollary 3.4.

On the other hand, in Figure 2, the pdf of $\widehat{a}_2$ has the Dickey-Fuller distribution. However, when $\mu = 0$, the pdf of $\widehat{a}_2$ is more symmetric by using the result of Phillips and Park (1988) as compared with that of $\widehat{a}_2^+$. As μ gets larger, the pdf of $\widehat{a}_2$ tends to skew more rightwards. This is because a larger value of μ implies a higher contemporaneous correlation between $u_{0,t}$ and $u_{2,t}$, hence the transformations of y_t and $\widehat{\Delta}_{0x}$ by (3.1) and (3.2) become more important. As seen from Figure 2, $\widehat{a}_2^+$ tends to be more efficient when μ becomes larger.

Figures 3 and 4 compare the distributions of $\widehat{a}_2^+$ and $\widehat{a}_2$ under various values of d and μ. The pdf's become flatter when d increases for all values of T and μ. This is consistent with the result of Theorem 3.1. The pdf's of $\widehat{a}_1^+$ are symmetric in Figure 3. When $T = 50$, $\widehat{a}_2^+$ has a upward bias and the size of bias increases with d. Nevertheless, the problem is eliminated as T increases in Figure 4(d).

Refer to Tables 1 - 4, MAD's of FM-SEA and OLS estimators of coefficients

a_1 and a_2 fall when the sample size increases. When the order of seasonal integration d increases, the MAD also increases. When the sample size is small, says $T = 50$, an increase in d induces a rise in Bias$_{avg}$'s for both $\widehat{a}_1^+$ and $\widehat{a}_2^+$. $\widehat{a}_2$ has the similar properties by Theorem 2.2. The results of $\widehat{a}_1^+$ and $\widehat{a}_2^+$ are caused by the performance of kernel estimates of long-run covariance matrices in finite samples. As μ increases, Bias$_{avg}$ of $\widehat{a}_2^+$ rises as compared with Figures 4 (a) and (b). Also, the RMSE$_{avg}$ increases with μ. However, as the samples size increases to 500, Bias$_{avg}$'s of $\widehat{a}_1^+$ and $\widehat{a}_2^+$ are greatly reduced. This says that the performance of the kernel estimates of long-run covariance matrices is improved by an increase in sample size T. Concerning with the t-statistic of $\widehat{a}_2^+$, an increase in μ will raise the Bias$_{avg}$ but lower the RMSE$_{avg}$. This implies that their variances become smaller. As a result, the t-statistic tends to reject the null hypothesis $a_2 = 0$. The RMSE$_{avg}$ of t-statistic for $\widehat{a}_2$ increases with μ unlike that for $\widehat{a}_2^+$. In addition, Bias$_{avg}$ of $\widehat{a}_2$ is larger than that of $\widehat{a}_2^+$. Thus, t-statistic for $\widehat{a}_2^+$ performs better than that for $\widehat{a}_2$. On the other hand, t-statistics for $\widehat{a}_1$ and $\widehat{a}_1^+$ do not change monotonically with μ. Nevertheless, the t-statistics for $\widehat{a}_1^+$ and $\widehat{a}_2^+$ have smaller bias than those of $\widehat{a}_1$ and $\widehat{a}_2$ except when $T = 50$. The reason is that the kernel estimates of long-run covariance matrices do not perform well in small samples.

Thus, the performance of $\widehat{a}_1^+$ and $\widehat{a}_2^+$ is fairly good under various values of d and μ especially for the non-stationary component. Of course, when T gets larger, the performance of FM-SEA estimator is improved.

6. Conclusion

In this paper, we have generalised the results of PFM by estimating the models with an unknown number of unit roots at various frequencies. The FM-SEA estiamtor for $I(0)$ component is $O(T^{1/2})$-consistent and has a normal asymptotic distribution. The asymptotic variance is different from that by Theorem 4.1(a) of PFM. The distribution for $SI(d,\theta)$ component is asymptotically mixed normal and varies with the order of seasonal integration. This means that the limiting distribution for $SI(d,\theta)$ component is proportional to the number of non-overlapping points on the unit circle. Furthermore, we introduce the deterministic trend component into the model and the result of $I(0)$ component is the same as that of Theorem 3.1 (a). In this case, the variates of stochastic trend and deterministic trend components are mixed. The $SI(d,\theta)$ component varies with the values of d, while the deterministic trend component is invariant to the values of d. Since the full coefficient matrix $\widehat{A}^{+}$ has a normal limit distribution, the conventional Wald statistic has a chi-squared limit distribution once the rank condition is satisfied.

On the other hand, we apply the technique of FM-SEA estimation to estimate a VAR(k) system. Our results are different from those in PFM because of the locations of complex unit roots. The distribution of stationary component, which includes stationary variables $\Delta_d x_{t-j}$ and cointegrated component $p_{1,t-1}$ in (4.6), is asymptotically normal. The distribution of non-cointegrated component is asymptotically mixed normal irrespective of the number of unit roots. In addition, we do not need to estimate the cointegrating vectors.

In finite samples, the FM-SEA estimator for non-stationary component performs very well under various degrees of contemporaneous correlation of $u_{0,t}$ and $u_{2,t}$ and the order of seasonal integration d. The average bias of FM-SEA estimator for $I(0)$ component is always smaller than those of OLS estimator. With respect to the $SI(d,\theta)$ component, the average bias of estimated coeffcient increases with d in small samples. When the sample size increases, the biases are reduced gradually. The t-statistics for FM-SEA estimated coefficients always have lower average biases and RMSEs especially when the sample size increases. This also indicates that the FM-SEA estimated coefficients have a better performance in testing hypothesis.

To conclude this paper, the FM-SEA estimation of models with an unknown number of seasonally integrated series yields a more general result of limiting distribution of full coefficient matrix. The number of unit roots is invariant to our results except in those extreme cases. Therefore, no unit root tests are required by this FM-SEA approach.

7. Mathematical Appendix

The lemmas given below are useful for proving the theorems and we will consider the case of untruncated kernels satisfying Assumption 2.2 (a) and (b). For the case of truncated kernels, the lemmas can be proved in the similar ways and the modifications have been discussed by PFM and Chang and Phillips (1995). Thus, we will not repeat those proofs for convenience.

The proofs given here follow closely to the proofs of Lemma 8.1 in PFM. The following facts are stated to help the proofs of lemmas and the details are discussed in PFM. Under the summability condition in Assumption 2.1 (a), we have

$$\sum_{j=0}^{\infty} j^a \left\|\Gamma(j)\right\| < \infty.$$

This fact is shown clearly in Phillips (1993, p.52) and this implies $\Gamma(K) = E(u_t u_{t-K}') = o(K^{-a})$. Also, $Var(\widehat{\Gamma}(K)) = O(T^{-1})$ as shown in Priestley (1970, p.326, equation 5.3.25) and Hannan (1970, p.212). Thus, the order of magnitude of $\widehat{\Gamma}(K)$ is given by

$$\widehat{\Gamma}(K) = O_p(T^{-1/2}) + o(K^{-a}) = o_p(1),$$

and this was shown in Priestley (1970, p.322). This relation also holds for $\widehat{\Gamma}(K-1)$

and $\widehat{\Gamma}(-K)$. Moreover, Assumptions 2.2 (b) and 2.3 (d) imply

$$w((K-1)/K) = O(K^{-2}) \text{ and } w((-K+1)/K) = O(K^{-2}).$$

Thus, the term $w((K-1)/K)\widehat{\Gamma}(K)$ appears in the Taylor's expansions of the kernel functions in the following proofs has the order of magnitude

$$w((K-1)/K)\widehat{\Gamma}(K) = O(K^{-2})o_p(1) = o_p(K^{-2}).$$

The approximation of kernel estimates by the Taylor's expansion is important to our proof because the sample autocovariance of the error terms with seasonally integrated processes is complicated and we can approximate the kernel estimates involving the autocovariance of $\Delta_d u_{1,t}$ and the other components of u_t by d $I(-1)$ processes. This approximation has been used in Chang and Phillips (1995) where the kernel estimate of $I(-2)$ process is approximated by two $I(-1)$ processes.

To simplify our notations in the following analysis, let $u_{b,t} = \left[\Delta_d u_{1,t}', u_{2,t}'\right]' = \Delta_d x_{b,t}' = D'\Delta_d x_t'$ and denote the long-run and one-sided long-run covariance matrices of $u_{0,t}$ and $u_{b,t}$ as Ω_{0b} and Δ_{0b} respectively. The long-run variance and one-sided long-run covariance matrices of $u_{b,t}$ are denoted by Ω_{bb} and Δ_{bb} respectively.

LEMMA 7.1. Under Assumptions 2.1, 2.2 and 2.3 (d), the followings hold :

(a) $\sum_{j=K_L}^{K_U} \Delta_d w(q(j)/K)\widehat{\Gamma}(j) = d \times K^{-2} w''(0) \sum_{j=-\infty}^{\infty} [q(j) - d/2]\, \Gamma(j)$

$+O_p(T^{-1/2}K^{-1/2})$,

(b) $\sum_{j=K_L}^{K_U} \Delta_d^2 w(q(j)/K)\widehat{\Gamma}(j) = d^2 K^{-2} w''(0)\Omega + O_p(T^{-1/2}K^{-3/2})$,

where $K_L = -K+d+1$, $K_U = K-d-1$, $K-1$ and $q(j) = j, d+d$.

PROOF. Define $B_* = \{j : |j| \leq K^*\}$ and $B^* = \{j : |j| > K^*,\ K_L \leq j \leq K_U\}$ for some $K^* = K^\delta$ with $\delta \in (0,1)$.

Part (a) For convenience, we choose $q(j) = j + d$ and rewrite the sum $\sum_{j=K_L}^{K_U} \Delta_d w((j+d)/K)\widehat{\Gamma}(j)$ as

$$\sum_{B_*} \Delta_d w((j+d)/K)\widehat{\Gamma}(j) + \sum_{B^*} \Delta_d w((j+d)/K)\widehat{\Gamma}(j). \tag{7.1}$$

Under Assumption 2.2, we can consider the approximation of $\Delta_d w((j+d)/K)$ by the Taylor's expansion.

$$\begin{aligned}
\Delta_d w((j+d)/K) &= w((j+d)/K) - w(j/K) \\
&= w'(j/K)(d/K) + \frac{1}{2} w''(j/K)(d/K)^2 + o(K^{-2}) \\
&= \left[w'(0) + w''(0)(j/K) + o(K^{-2})\right](d/K) \\
&\quad + \frac{1}{2}\left[w''(0) + o(1)\right](d^2/K^2) + o(K^{-2}) \\
&= w''(0)(dj/K^2) + w''(0)((d^2/2)/K^2) + o(K^2) \\
&= d \times K^{-2} w''(0)(j + d/2)(1 + o(1)).
\end{aligned}$$

Then, the first sum of (7.1) becomes

$$\sum_{|j|\leq K^*} \Delta_d w((j+d)/K)\widehat{\Gamma}(j) = d \times K^{-2} w''(0) \left[\sum_{|j|\leq K^*} (j+d/2)\widehat{\Gamma}(j)(1+o(1)) \right].$$

The mean of the term with squared bracket is given by

$$\sum_{|j|\leq K^*} (j+d/2)(1-|j|/T)\Gamma(j) \rightarrow \sum_{j=-\infty}^{\infty} (j+d/2)\Gamma(j).$$

Note that mean used here is provided by Priestly (1981, vol. 1, equation 5.3.14). Next, we consider the second sum of (7.1) and approximate the kernel function by the Taylor's expansion

$$\Delta_d w((j+d)/K) = d \times K^{-1} w'(\theta_j),$$

for some $\theta_j \in (j/K, (j+d)/K)$. Then, the second sum of (7.1) can be rewritten as

$$d \times K^{-1} \sum_{B^*} w'(\theta_j)\widehat{\Gamma}(j),$$

whose mean is given by

$$d \times K^{-1} \sum_{B^*} w'(\theta_j)(1-|j|/T)\Gamma(j).$$

The modulus of the mean above is dominated by

$$d \times K^{-1} \left[\sup_{K_L \leq j \leq K_U} |w'(\theta_j)| \right] \sum_{|j|>K^*} \|\Gamma(j)\|$$

$$
\begin{aligned}
&\leq d \times K^{-1} M \sum_{|j|>K^*} \|C_s\| \, \|C_{s+j}\| \\
&\leq d \times K^{-1} K^{*-a} M \sum_{|j|>K^*} \sum_{s=0}^{\infty} (s+j)^a \|C_s\| \, \|C_{s+j}\| \\
&\leq d \times K^{-1} K^{-a\delta} M \sum_{s=0}^{\infty} \|C_s\| \sum_{r=0}^{\infty} r^a \|C_r\|,
\end{aligned}
$$

where $w'(.)$ is uniformly bounded under Assumption 2.2, that is, $\exists M$ such that $|w'(x)| \leq M \; \forall x \in R$. Note that we can choose δ so that $a\delta > 1$. Then, the mean of second sum of (7.1) has the order of $o(K^{-2})$ as $K \to \infty$. Hence, the first sum of (7.1) dominates the second sum.

Now, we consider the variance matrix of the sum in (7.1) by writing it as

$$
\sum_{j=K_L}^{K_U} \Delta_d w((j+d)/K)\widehat{\Gamma}(j) = d \times K^{-1} \sum_{j=K_L}^{K_U} w'(j/K)\widehat{\Gamma}(j)(1+o(1)).
$$

Then, by Hannan (1970, Theorem 9), we have

$$
\begin{aligned}
&\lim_{T\to\infty} TKVar\left[vec\left(d \times K^{-1} \sum_{j=K_L}^{K_U} w'(j/K)\widehat{\Gamma}(j)\right)\right] \\
&= d^2 \lim_{T\to\infty} TK^{-1}Var\left[vec\left(\sum_{j=K_L}^{K_U} w'(j/K)\widehat{\Gamma}(j)\right)\right] \\
&= \text{constant}.
\end{aligned}
$$

This implies that the variance of the dominant term of (7.1) is $O(T^{-1}K^{-1})$.

The preceding obtained results lead us to deduce that

$$
\sum_{j=K_L}^{K_U} \Delta_d w((j+d)/K)\widehat{\Gamma}(j) = d \times K^{-2} \sum_{j=-\infty}^{\infty} (j+d/2)\Gamma(j) + O_p(T^{-1/2}K^{-1/2}).
$$

This is the required result. Note that the same analysis can be applied for $q(j) = j$. ■

Part (b) As in the proof of part (a), we choose $q(j) = j + d$ and rewrite the sum $\sum_{j=K_L}^{K_U} \Delta_d^2 w(q(j)/K)\widehat{\Gamma}(j)$ as

$$\sum_{B_*} \Delta_d^2 w((j+d)/K)\widehat{\Gamma}(j) + \sum_{B^*} \Delta_d^2 w((j+d)/K)\widehat{\Gamma}(j). \tag{7.2}$$

Under Assumption 2.2, we can consider the approximation of $\Delta_d^2 w((j+d)/K)$ by the second-order Taylor's expansion.

$$\begin{aligned}
\Delta_d^2 w((j+d)/K) &= \Delta_d \left[w((j+d)/K) - w(j/K)\right] \\
&= [w((j+d)/K) - w(j/K)] + [w((j-d)/K) - w(j/K)] \\
&= w'(j/K)(d/K) + (1/2)w''(j/K)(d/K^2) \\
&\quad - w'(j/K)(d/K) + (1/2)w''(j/K)(d/K)^2 + o(K^{-2}) \\
&= d^2 K^{-2} w''(0)(1+o(1)) + o(K^{-2}) \\
&= d^2 K^{-2} w''(0)(1+o(1)).
\end{aligned}$$

Since $w''(.)$ is continuous and uniformly bounded under Assumption 2.2 (b), then the first sum of (7.2) has the mean

$$d^2 K^{-2} w''(0) \sum_{|j| \leq K^*} (1 - |j|/T)\Gamma(j).$$

The mean scaled by K^2 becomes

$$d^2 w''(0) \sum_{|j| \leq K^*} (1 - |j|/T)\Gamma(j) \to d^2 w''(0)\Omega.$$

The second sum of (7.2) is approximated by the first-order Taylor's expansion

$$\sum_{B^*} \Delta_d^2 w((j+d)/K)\widehat{\Gamma}(j) = d^2 K^{-2} \sum_{B^*} w''(\theta_j)\widehat{\Gamma}(j),$$

where $\theta_j \in (j/K, (j+d)/K)$ and it has the mean

$$d^2 K^{-2} \sum_{B^*} w''(\theta_j)(1 - |j|/T)\Gamma(j),$$

and the modulus of the second sum of (7.2) is dominated by

$$d^2 K^{-2} \left[\sup_{K_L \leq j \leq K_U} |w'(\theta_j)| \right] \sum_{|j| > K^*} \|\Gamma(j)\| \leq O(K^{-2-a\delta}),$$

where $\delta \in (0, 1)$ and this is followed directly from the proof of part (a). Also, the variance matrix of (7.2) is considered by rewriting (7.2) as

$$\sum_{j=K_L}^{K_U} \Delta_d^2 w((j+d)/K)\widehat{\Gamma}(j) = d^2 K^{-2} \sum_{j=K_L}^{K_U} w''(j/K)\widehat{\Gamma}(j)(1 + o(1)).$$

As in the proof of part (a), the variance matrix is $O(T^{-1}K^{-3})$ because

$$\lim_{T \to \infty} TK^3 Var \left[vec \left(d^2 K^{-2} \sum_{j=K_L}^{K_U} w''(j/K)\widehat{\Gamma}(j) \right) \right]$$

$$
\begin{aligned}
&= d^4 \lim_{T\to\infty} TK^{-1} Var\left[vec\left(\sum_{j=K_L}^{K_U} w''(j/K)\widehat{\Gamma}(j) \right) \right] \\
&= \text{constant.}
\end{aligned}
$$

Thus, we obtain the required result. ■

LEMMA 7.2. Under Assumptions 2.1, 2.2 and 2.3 (d), the followings hold :

(a) $\widehat{\Omega}_{\Delta_d u_1 \Delta_d u_1} = -d^2 K^{-2} w''(0)\Omega_{11} + O_p(T^{-1/2}K^{-3/2}) + o_p(K^{-2})$,

(b) $\widehat{\Omega}_{u_0 \Delta_d u_1} = d \times K^{-2} w''(0)\phi_{01} + O_p(T^{-1/2}K^{-1/2}) + o_p(K^{-2})$,

(c) $\widehat{\Omega}_{u_2 \Delta_d u_1} = d \times K^{-2} w''(0)\phi_{21} + O_p(T^{-1/2}K^{-1/2}) + o_p(K^{-2})$,

(d) $\widehat{\Omega}_{\Delta_d u_1 u_2} = d \times K^{-2} w''(0)\phi_{12} + O_p(T^{-1/2}K^{-1/2}) + o_p(K^{-2})$,

(e) $\widehat{\Omega}_{0\Delta_d u_1} := \widehat{\Omega}_{\widehat{u}_0 \Delta_d u_1} = \widehat{\Omega}_{u_0 \Delta_d u_1} + O_p(T^{-1}) + O_P(T^{-1/2}K^{-2})$,

(f)
$$
\begin{aligned}
\widehat{\Omega}_{0b}\widehat{\Omega}_{bb}^{-1} = \Big[&-d^{-1}\left(\phi_{01} - \Omega_{02}\Omega_{22}^{-1}\phi_{21}\right)\Omega_{11}^{-1} + O_p(T^{-1/2}K^{3/2}) + o_p(T^{-1/2}K^{3/2}) \\
&\vdots\ \Omega_{02}\Omega_{22}^{-1} + o_p(1)\Big],
\end{aligned}
$$

where $\phi_{01} = \sum_{j=-\infty}^{\infty}(j - d/2)\Gamma_{u_0 u_1}(j)$, $\phi_{21} = \sum_{j=-\infty}^{\infty}(j - d/2)\Gamma_{u_2 u_1}(j)$ and $\phi_{12} = \sum_{j=-\infty}^{\infty}(j - d/2)\Gamma_{u_1 u_2}(j)$.

PROOF.

Part (a) Recall from (2.13), the kernel estimate $\widehat{\Omega}$ is decomposed into

$$
\widehat{\Gamma}_{\Delta_d u_1 \Delta_d u_1}(j) \;=\; T^{-1} \sum_{t=j+1}^{T} \Delta_d u_{1,t} \Delta_d u_{1,t-j}'
$$

$$
\begin{aligned}
&= T^{-1}\sum_{t=j+1}^{T}\left[u_{1,t}-u_{1,t-d}\right]\left[u_{1,t-j}-u_{1,t-d-j}\right]' \\
&= T^{-1}\sum_{t=j+1}^{T}\Big[u_{1,t}u'_{1,t-j}-u_{1,t-d}u'_{1,t-j}-u_{1,t}u'_{1,t-d-j}+ \\
&\quad +u_{1,t-d}u'_{1,t-d-j}\Big] \\
&= \widehat{\Gamma}_{u_1u_1}(j)-\widehat{\Gamma}_{u_1u_1}(j-d)-\widehat{\Gamma}_{u_1u_1}(j+d)+\widehat{\Gamma}_{u_1u_1}(j) \\
&= -\Delta_d^2\widehat{\Gamma}_{u_1u_1}(j+d).
\end{aligned}
$$

Then, it follows that

$$
\begin{aligned}
\widehat{\Omega}_{\Delta_d u_1\Delta_d u_1} &= -\sum_{j=-K+1}^{K-1} w(j/K)\Delta_d^2\widehat{\Gamma}_{u_1u_1}(j+d) \\
&= -\sum_{j=-K+1}^{K-1} w(j/K)\left[-2\widehat{\Gamma}_{u_1u_1}(j)+\widehat{\Gamma}_{u_1u_1}(j-d)+\widehat{\Gamma}_{u_1u_1}(j+d)\right] \\
&= -\Bigg[-2\sum_{j=-K+1}^{K-1} w(j/K)\widehat{\Gamma}_{u_1u_1}(j)+\sum_{j=-K-d+1}^{K-d-1} w((j+d)/K)\widehat{\Gamma}_{u_1u_1}(j) \\
&\quad +\sum_{j=-K+d+1}^{K+d-1} w((j-d)/K)\widehat{\Gamma}_{u_1u_1}(j)\Bigg] \\
&= -\sum_{j=-K+d+1}^{K-d-1}\left[w((j+d)/K)-2w(j/K)+w((j-d)/K)\right]\widehat{\Gamma}_{u_1u_1}(j) \\
&\quad +\eta \\
&= -\sum_{j=-K+d+1}^{K-d-1}\Delta_d^2 w((j+d)/K)\widehat{\Gamma}_{u_1u_1}(j)+o_p(K^{-2}),
\end{aligned}
$$

where

$$
\eta = 2\sum_{j=K-d}^{K-1} w(j/K)\widehat{\Gamma}_{u_1u_1}(j)+2\sum_{j=K+1}^{-K+d} w(j/K)\widehat{\Gamma}_{u_1u_1}(j)
$$

$$
\begin{aligned}
& - \sum_{j=-K-d+1}^{-K+d} w((j+d)/K)\widehat{\Gamma}_{u_1u_1}(j) - \sum_{j=K-d}^{K+d-1} w((j-d)/K)\widehat{\Gamma}_{u_1u_1}(j) \\
= \; & o_p(K^{-2}).
\end{aligned}
$$

Together with Lemma 7.1 (b), we have

$$
\widehat{\Omega}_{\Delta_d u_1 \Delta_d u_1} = -d^2 K^{-2} w''(0)\Omega_{11} + O_p(T^{-1/2}K^{-3/2}) + o_p(K^{-2}). \qquad \blacksquare
$$

Part (b) Consider the expression of kernel estimate $\widehat{\Omega}_{u_0\Delta_d u_1}$ directly.

$$
\begin{aligned}
\widehat{\Omega}_{u_0\Delta_d u_1} &= \sum_{j=-K+1}^{K-1} w(j/K)\widehat{\Gamma}_{u_0\Delta_d u_1}(j) \\
&= \sum_{j=-K+1}^{K-1} w(j/K)\left[\widehat{\Gamma}_{u_0u_1}(j) - \widehat{\Gamma}_{u_0u_1}(j+d)\right] \\
&= \sum_{j=-K+1}^{K-1} w(j/K)\widehat{\Gamma}_{u_0u_1}(j) - \sum_{j=-K+d+1}^{K+d-1} w((j-d)/K)\widehat{\Gamma}_{u_0u_1}(j) \\
&= \sum_{j=-K+d+1}^{K-1} \left[w(j/K) - w((j-d)/K)\right]\widehat{\Gamma}_{u_0u_1}(j) \\
&\quad + \sum_{j=-K+1}^{-K+d} w(j/K)\widehat{\Gamma}_{u_0u_1}(j) - \sum_{j=K}^{K+d-1} w((j-d)/K)\widehat{\Gamma}_{u_0u_1}(j) \\
&= \sum_{j=-K+d+1}^{K-1} \Delta_d w(j/K)\widehat{\Gamma}_{u_0u_1}(j) + o_p(K^{-2}).
\end{aligned}
$$

Thus, $\widehat{\Omega}_{u_0\Delta_d u_1} = d \times K^{-2}w''(0)\phi_{01} + O_p(T^{-1/2}K^{-1/2}) + o_p(K^{-2})$ by Lemma 7.1 (a). $\blacksquare$

Part (c) and (d) can be proved by the same way as in part (b), so we will not repeat the proof here.

Part (e) We write

$$
\begin{aligned}
&\widehat{\Omega}_{\widehat{u}_0\Delta_d u_1} \\
=\; &\widehat{\Omega}_{u_0\Delta_d u_1} - \sum_{j=-K+1}^{K-1} w(j/K)(\widehat{A}-A)\widehat{\Gamma}_{x\Delta_d u_1}(j) \\
=\; &\widehat{\Omega}_{u_0\Delta_d u_1} - \sum_{j=-K+1}^{-K+d} w(j/K)(\widehat{A}-A)\widehat{\Gamma}_{xu_1}(j) + \sum_{j=K}^{K+d-1} w(j/K)(\widehat{A}-A)\widehat{\Gamma}_{xu_1}(j) \\
&- \sum_{j=-K+d+1}^{K+d-1} \Delta_d w(j/K)(\widehat{A}-A)\widehat{\Gamma}_{xu_1}(j). \qquad (7.3)
\end{aligned}
$$

Then,

$$
\begin{aligned}
&\widehat{\Gamma}_{xu_1}(j) \\
=\; &D\left[\widehat{\Gamma}_{x_1u_1}(j) + \widehat{\Gamma}_{x_2u_1}(j)\right] \\
=\; &D\left[\widehat{\Gamma}_{u_1u_1}(j) + \widehat{\Gamma}_{x_2u_1}(j)\right] \\
=\; &O_p(1) \qquad \text{for } -K+1 \le j \le -K+d \text{ and } K \le j \le K+d-1,
\end{aligned}
$$

because the first term is the autocovariance of stationary error component $u_{1,t}$ while the second term converges in distribution to some stochastic integral with bounded variation as shown in Phillips (1988b). Also, using the fact that $w(j/K) = O(K^{-2})$ for $-K+1 \le j \le -K+d$ and $K \le j \le K+d-1$, $\widehat{A} - A = O_p(T^{-1/2})$, the second and third terms in (7.3) are $O_p(T^{-1/2}K^{-2})$.

Now, we consider the fourth term in the expression of (7.3)

$$\sum_{j=-K+d+1}^{K+d-1} \Delta_d w(j/K)\left(\widehat{A}-A\right)\widehat{\Gamma}_{xu_1}(j)$$

$$= \sum_{j=-K+d+1}^{K+d-1} \Delta_d w(j/K)\left(\widehat{A}_1-A_1\right)\widehat{\Gamma}_{x_1u_1}(j)$$

$$+ \sum_{j=-K+d+1}^{K+d-1} \Delta_d w(j/K)\left(\widehat{A}_2-A_2\right)\widehat{\Gamma}_{x_2u_1}(j). \tag{7.4}$$

Using the fact that $\widehat{A}_1 - A_1 = O_p(T^{-1/2})$ and the result of part (b), we have

$$\sum_{j=-K+d+1}^{K+d-1} \Delta_d w(j/K)\left(\widehat{A}_1-A_1\right)\widehat{\Gamma}_{x_1u_1}(j)$$

$$= \left[O_p(K^{-2}) + O_p(T^{-1/2}K^{-1/2})\right] O_p(T^{-1/2})$$

$$= O_p(T^{-1/2}K^{-2}) + O_p(T^{-1}K^{-1/2}).$$

For the second sum of (7.4), we see that $\widehat{A}_2 - A_2 = O_p(T^{-1})$ and $\sum_{j=-K+d+1}^{K+d-1} \Delta_d w(j/K)\widehat{\Gamma}_{x_2u_1}(j) = K^{-1}\sum_{j=-K+d+1}^{K+d-1} w'(\theta_j/K)\widehat{\Gamma}_{x_2u_1}(j) = O_p(1)$ as in the proof of Theorem 3.1 of Phillips (1991b, p.432-433) where $\theta_j \in ((j-d)/K, j/K)$. This deduces the second sum of (7.4) is $O_p(T^{-1})$. ■

Part (f) By partitioned matrix inversion of $\widehat{\Omega}_{bb}$, we have

$$\widehat{\Omega}_{bb}^{-1} = \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix},$$

where $B_{11} = \widehat{\Omega}_{\Delta_d u_1 \Delta_d u_1} - \widehat{\Omega}^{-1}_{\Delta_d u_1 \Delta_d u_1} \widehat{\Omega}_{\Delta_d u_1 u_2} \widehat{\Omega}^{-1}_{u_2 u_2 . \Delta_d u_1} \widehat{\Omega}_{u_2 \Delta_d u_1} \widehat{\Omega}^{-1}_{\Delta_d u_1 \Delta_d u_1}$, $B_{12} = -\widehat{\Omega}^{-1}_{\Delta_d u_1 \Delta_d u_1} \widehat{\Omega}_{\Delta_d u_1 u_2} \widehat{\Omega}^{-1}_{u_2 u_2 . \Delta_d u_1}$, $B_{21} = -\widehat{\Omega}^{-1}_{u_2 u_2 . \Delta_d u_1} \widehat{\Omega}_{u_2 \Delta_d u_1} \widehat{\Omega}^{-1}_{\Delta_d u_1 \Delta_d u_1}$ and $B_{22} = \widehat{\Omega}^{-1}_{u_2 u_2 . \Delta_d u_1} = \left[\widehat{\Omega}_{u_2 u_2} - \widehat{\Omega}_{u_2 \Delta_d u_1} \widehat{\Omega}^{-1}_{\Delta_d u_1 \Delta_d u_1} \widehat{\Omega}_{\Delta_d u_1 u_2} \right]^{-1}$.

Then,

$$
\begin{aligned}
\widehat{\Omega}_{0b} \widehat{\Omega}^{-1}_{bb} &= \widehat{\Omega}_{0 \Delta_d u_1} \left[\begin{array}{ccc} B_{11} & \vdots & B_{12} \end{array} \right] + \widehat{\Omega}_{0 u_2} \left[\begin{array}{ccc} B_{21} & \vdots & B_{22} \end{array} \right] \\
&= \left[\begin{array}{ccc} X_{01} & \vdots & X_{02} \end{array} \right].
\end{aligned}
$$

Using Lemma 7.2 (a), (c) and (d), we have

$$
\begin{aligned}
\widehat{\Omega}_{u_2 u_2 . \Delta_d u_1} &= \widehat{\Omega}_{u_2 u_2} + O_p(K^{-2}) + O_p(T^{-1} K) + O_p(T^{-1/2} K^{-1/2}) \\
&= \widehat{\Omega}_{u_2 u_2} + o_p(1) \to_p \Omega_{22}.
\end{aligned}
$$

Also, using Lemma 7.2 (a), (b), (c), (d), (e) and the above result, we have

$$
\begin{aligned}
X_{01} &= \left[-d^{-1} \phi_{01} + O_p(T^{-1/2} K^{3/2}) \right] \Omega_{11}^{-1} - \left[-d^{-1} \phi_{01} + O_p(T^{-1/2} K^{3/2}) \right] \Omega_{11}^{-1} \\
&\quad \times \left[O_p(K^{-2}) + O_p(T^{-1/2} K^{-1/2}) \right] \left[\Omega_{22} + o_p(1) \right]^{-1} \\
&\quad \times \left[-d^{-1} \phi_{21} + O_p(T^{-1/2} K^{3/2}) \right] \Omega_{11}^{-1} \\
&\quad - \left[\Omega_{02} + o_p(1) \right] \left[\Omega_{22} + o_p(1) \right]^{-1} \left[-d^{-1} \phi_{21} + O_p(T^{-1/2} K^{3/2}) \right] \Omega_{11}^{-1} \\
&= d^{-1} \left[\phi_{01} - \Omega_{02} \Omega_{22}^{-1} \phi_{21} \right] \Omega_{11}^{-1} + O_p(T^{-1/2} K^{3/2}) + o_p(T^{-1/2} K^{3/2}). \\
X_{02} &= - \left[-d^{-1} \phi_{01} + O_p(T^{-1/2} K^{3/2}) \right] \left[\Omega_{11} + o_p(1) \right]^{-1}
\end{aligned}
$$

$$\times \left[O_p(K^{-2}) + O_p(T^{-1/2}K^{-1/2})\right] \Omega_{22}^{-1} + \Omega_{02}\Omega_{22}^{-1} + o_p(1)$$

$$= \Omega_{02}\Omega_{22}^{-1} + o_p(1). \qquad \blacksquare$$

LEMMA 7.3. Under Assumptions 2.1, 2.2 and 2.3 (d), the followings hold :

(a) $\left[T^{-1}\Delta_d U_1' S(U_1, d) - \widehat{\Delta}_{\Delta_d u_1 \Delta_d u_1}\right] = d^2 K^{-2} w''(0)$

$\times \left[\Delta_{11} - 2^{-1}\Sigma_{11} + d^{-1}\sum_{j=1}^{d-1}(j - d/2)\Gamma_{u_1u_1}(j)\right] + O_p(T^{-1/2}K^{-1}) + o(K^2),$

(b) $T^{-1}U_2' S(U_1, d) - \widehat{\Delta}_{u_2\Delta_d u_1} = -d \times K^{-2} w''(0)\psi_{21} + O_p(T^{-1/2}K^{-1/2}) + o_p(K^{-2}),$

(c) $T^{-1}\Delta_d U_1' S(X_2, d) - \widehat{\Delta}_{\Delta_d u_1 u_2} = -d \times K^{-2} w''(0)\psi_{12} + O_p(T^{-1/2}) + o_p(K^{-2}),$

(d) $\widehat{\Delta}_{0\Delta_d u_1} := \widehat{\Delta}_{\widehat{u}_0 \Delta_d u_1} = \sum_{j=1}^{d-1}\widehat{\Gamma}_{u_0u_1}(j) + O_p(T^{-1}) + O_p(T^{-1/2}K^{-1/2}),$

(e) $\widehat{\Delta}_{0u_2} := \widehat{\Delta}_{\widehat{u}_0 u_2} = \Delta_{02} + O_p(T^{-1/2}K^{1/2}),$

where $\psi_{21} = \sum_{j=d}^{\infty}(j - d/2)\Gamma_{u_1u_2}(j)$ and $\psi_{12} = \sum_{j=d}^{\infty}(j + d/2)\Gamma_{u_2u_1}(j)$.

PROOF. Define $B_* = \{j : d \le j \le K^*\}$ and $B^* = \{j : K^* < j \le K + d - 1\}$ for some $K^* = K^\delta$ with $\delta \in (0, 1)$.

Part (a) We observe that

$$\begin{aligned}
& T^{-1}\Delta_d U_1' S(U_1, d) - \widehat{\Delta}_{\Delta_d u_1 \Delta_d u_1} \\
= \; & T^{-1}\Delta_d U_1' S(U_1, d) - \sum_{j=0}^{K-1} w(j/K)\left[\widehat{\Gamma}_{\Delta_d u_1 u_1}(j) - \widehat{\Gamma}_{\Delta_d u_1 u_1}(j+d)\right] \\
= \; & T^{-1}\Delta_d U_1' S(U_1, d) - \sum_{j=d}^{K+d-1}\left[w(j/K) - w((j-d)/K)\right]\widehat{\Gamma}_{\Delta_d u_1 u_1}(j)
\end{aligned}$$

$$
\begin{aligned}
& + \sum_{j=K}^{K+d-1} w(j/K)\widehat{\Gamma}_{\Delta_d u_1 u_1}(j) - \sum_{j=0}^{d-1} w(j/K)\widehat{\Gamma}_{\Delta_d u_1 u_1}(j) \\
= \;& T^{-1}\Delta_d U_1' S(U_1, d) - \sum_{j=d}^{K+d-1} \left[w(j/K) - w((j-d)/K)\right] \widehat{\Gamma}_{\Delta_d u_1 u_1}(j) \\
& + \sum_{j=K}^{K+d-1} w(j/K)\widehat{\Gamma}_{\Delta_d u_1 u_1}(j) - \sum_{j=0}^{d-1} \left[w(j/K) - w(0)\right] \widehat{\Gamma}_{\Delta_d u_1 u_1}(j) \\
& - \sum_{j=0}^{d-1} \widehat{\Gamma}_{\Delta_d u_1 u_1}(j) \\
= \;& - \sum_{j=d}^{K+d-1} \Delta_d w(j/K)\widehat{\Gamma}_{\Delta_d u_1 u_1}(j) + O_p(T^{-1/2}K^{-1}) + o_p(K^{-2}),
\end{aligned}
$$

since $\sum_{j=0}^{d-1} \left[w(j/K) - w(0)\right] \widehat{\Gamma}_{\Delta_d u_1 u_1}(j) = K^{-1} \sum_{j=0}^{d-1} w'(\gamma_j)\widehat{\Gamma}_{\Delta_d u_1 u_1}(j)$ $= O_p(T^{-1/2}K^{-1})$ where $\gamma_j \in (0, j/K)$.

Now, we consider the first term of the above expression as in the proof of Lemma 7.1(a) and write it as

$$
- \sum_{B_*} \Delta_d w(j/K)\widehat{\Gamma}_{\Delta_d u_1 u_1}(j) - \sum_{B^*} \Delta_d w(j/K)\widehat{\Gamma}_{\Delta_d u_1 u_1}(j), \tag{7.5}
$$

Under Assumption 2.2, the term $\Delta_d w(j/K)$ can be approximated by the Taylor's expansion, we have

$$
d \times K^{-2}(j - d/2)w''(0)(1 + o(1)).
$$

Then, the first term of (7.5) has the mean

$$
\begin{aligned}
& -d \times K^{-2}w''(0) \sum_{j=d}^{K^*} (j - d/2)(1 - |j|/T)\Gamma_{\Delta_d u_1 u_1}(j) \\
= \;& -d \times K^{-2}w''(0) \sum_{j=d}^{K^*} \left[(j - d/2)\Gamma_{u_1 u_1}(j) - (j - d/2)\Gamma_{u_1 u_1}(j - d)\right] + o(K^{-2})
\end{aligned}
$$

$$
\begin{aligned}
&= -d \times K^{-2} w''(0) \left[\sum_{j=d}^{K^*} (j - d/2) \Gamma_{u_1 u_1}(j) - \sum_{j=0}^{K^*-d} (j + d/2) \Gamma_{u_1 u_1}(j) \right] + o(K^{-2}) \\
&= -d \times K^{-2} w''(0) \left[\sum_{j=1}^{K^*-d} (j - d/2 - j - d/2) \Gamma_{u_1 u_1}(j) - \sum_{j=1}^{d-1} (j - d/2) \Gamma_{u_1 u_1}(j) \right. \\
&\quad \left. + \sum_{j=K^*-d+1}^{K^*} (j - d/2) \Gamma_{u_1 u_1}(j) - (d/2) \Gamma_{u_1 u_1}(0) \right] + o(K^{-2}) \\
&= d^2 K^{-2} w''(0) \left[\sum_{j=1}^{K^*-d} \Gamma_{u_1 u_1}(j) + 2^{-1} \Gamma_{u_1 u_1}(0) + d^{-1} \sum_{j=1}^{d-1} (j - d/2) \Gamma_{u_1 u_1}(j) \right] \\
&\quad + o(K^{-2}).
\end{aligned}
$$

Thus,

$$
\begin{aligned}
& K^2 E \left[d \times K^{-2} w''(0) \sum_{j=d}^{K^*} (j - d/2) \widehat{\Gamma}_{\Delta_d u_1 u_1}(j) \right] \\
\rightarrow_p \quad & d^2 w''(0) \left[\Delta_{11} + 2^{-1} \Sigma_{11} + d^{-1} \sum_{j=1}^{d-1} (j - d/2) \Gamma_{u_1 u_1}(j) \right] \\
= \quad & d^2 w''(0) \left[\Delta_{11} + 2^{-1} \Sigma_{11} + d^{-1} \sum_{j=1}^{d-1} (j - d/2) \Gamma_{u_1 u_1}(j) \right]
\end{aligned}
$$

Next, consider the mean of the second term of (7.5)

$$
-d \times K^{-1} \sum_{j=K^*+1}^{K+d-1} w'(\theta_j)(1 - |j|/T) \Gamma_{\Delta_d u_1 u_1}(j)(1 + O(K^{-1})),
$$

as $\Delta_d w(j/K)$ is approximated by the Taylor's expansion for $\theta_j \in ((j-d)/K, j/K)$.

As in the proof of Lemma 7.1 (a), the modulus of this expression is dominated by

$$
-d \times K^{-1} \left[\sup_{d \le j \le K+d-1} |w'(\theta_j)| \right] \sum_{j=K^*+1}^{K+d-1} \|\Gamma_{\Delta_d u_1 u_1}(j)\| (1 + O(K^{-1})) = o(K^{-2}),
$$

and the variance matrix of (7.5) is $O(T^{-1}K^{-3})$. ■

Part (b)

$$
\begin{aligned}
& T^{-1}U_2'S(U_1,d) - \widehat{\Delta}_{u_2\Delta_d u_1} \\
= \; & T^{-1}U_2'S(U_1,d) - \sum_{j=0}^{K-1} w(j/K)\left[\widehat{\Gamma}_{u_2u_1}(j) - \widehat{\Gamma}_{u_2u_1}(j+d)\right] \\
= \; & T^{-1}U_2'S(U_1,d) - \sum_{j=d}^{K+d-1}\left[w(j/K) - w((j-d)/K)\right]\widehat{\Gamma}_{u_2u_1}(j) \\
& - \sum_{j=0}^{d-1}\left[w(j/K) - w(0)\right]\widehat{\Gamma}_{u_2u_1}(j) - \sum_{j=0}^{d-1}\widehat{\Gamma}_{u_2u_1}(j) + o_p(K^{-2}) \\
= \; & -\sum_{j=d}^{K+d-1}\Delta_d w(j/K)\widehat{\Gamma}_{u_2u_1}(j) + O_p(T^{-1/2}K^{-1}) + o_p(K^{-2}),
\end{aligned}
$$

by the same tricks as in the proof of part (a).

Now, we rewrite the first term of the above expression as

$$
-\sum_{B_*}\Delta_d w(j/K)\widehat{\Gamma}_{u_2u_1}(j) - \sum_{B^*}\Delta_d w(j/K)\widehat{\Gamma}_{u_2u_1}(j). \tag{7.6}
$$

Under Assumption 2.2, the term $\Delta_d w(j/K)$ can be approximated by $d\times K^{-2}(j-d/2)w''(0)(1+o(1))$ and the mean of the first sum of (7.6) is given by

$$
\begin{aligned}
& -d\times K^{-2}w''(0)\sum_{j=d}^{K^*}(j-d/2)(1-|j|/T)\Gamma_{u_2u_1}(j)(1+o(1)) \\
\to \; & -d\times K^{-2}w''(0)\sum_{j=d}^{\infty}(j-d/2)\Gamma_{u_2u_1}(j) \\
= \; & -d\times K^{-2}w''(0)\psi_{21}.
\end{aligned}
$$

The analysis of the second sum of (7.6) is the same as that in the proof of part (a) and the variance matrix of (7.4) is $O(T^{-1}K^{-1})$. ■

Part (c) To prove this lemma, we consider

$$
\begin{aligned}
& T^{-1}\Delta_d U_1' S(X_2,d) - \widehat{\Delta}_{\Delta_d u_1 u_2} \\
= \;& T^{-1}(U_1' - U_{1,-d}')S(X_2,d) - \widehat{\Delta}_{\Delta_d u_1 u_2} \\
= \;& T^{-1}S(u_{1,T}x_{2,T}', d) - T^{-1}U_{1,-d}'S(U_2,d) - \widehat{\Delta}_{\Delta_d u_1 u_2}.
\end{aligned}
$$

Next, we observe that

$$
\begin{aligned}
& \widehat{\Delta}_{\Delta_d u_1 u_2} + T^{-1}U_{1,-d}'S(U_2,d) \\
= \;& \sum_{j=0}^{K-1} w(j/K)\widehat{\Gamma}_{\Delta_d u_1 u_2}(j) + T^{-1}U_{1,-d}'S(U_2,d) \\
= \;& \sum_{j=0}^{K-1} w(j/K)\left[\widehat{\Gamma}_{u_1u_2}(j) - \widehat{\Gamma}_{u_1u_2}(j-d)\right] + T^{-1}U_{1,-d}'S(U_2,d) \\
= \;& \sum_{j=0}^{K-d-1} \left[w(j/K) - w((j+d)/K)\right]\widehat{\Gamma}_{u_1u_2}(j) + T^{-1}U_{1,-d}'S(U_2,d) \\
& + \sum_{j=K-d}^{K-1} w(j/K)\widehat{\Gamma}_{u_1u_2}(j) - \sum_{j=-d}^{-1} w((j+d)/K)\widehat{\Gamma}_{u_1u_2}(j) \\
= \;& -\sum_{j=0}^{K-d-1} \Delta_d w((j+d)/K)\widehat{\Gamma}_{u_1u_2}(j) + T^{-1}U_{1,-d}'S(U_2,d) \\
& + \sum_{j=K-d}^{K-1} w(j/K)\widehat{\Gamma}_{u_1u_2}(j) - \sum_{j=-d}^{-1} \left[w((j+d)/K) - w(0)\right]\widehat{\Gamma}_{u_1u_2}(j) \\
& - \sum_{j=-d}^{-1} \widehat{\Gamma}_{u_1u_2}(j) \\
= \;& -\sum_{j=0}^{K-d-1} \Delta_d w((j+d)/K)\widehat{\Gamma}_{u_1u_2}(j) + O_p(T^{-1/2}) + o_p(K^{-2}).
\end{aligned}
$$

where the second and fifth terms are $O_p(T^{-1/2})$ and the fourth term is by the tricks as in the proof of part (a).

The proof can be followed by part (b) and the mean of the first term is given by

$$
\begin{aligned}
& -d \times K^{-2} \sum_{j=0}^{K-d-1} w''(0)\,(j+d/2)(1-|j|/T)\Gamma_{u_1 u_2}(j) \\
\rightarrow \quad & -d \times K^{-2} w''(0) \sum_{j=0}^{\infty} (j+d/2)\Gamma_{u_1 u_2}(j) \\
= \quad & -d \times K^{-2} w''(0)\psi_{12}.
\end{aligned}
$$

Also, the variance matrix of the sum is $O(T^{-1}K^{-1})$. ■

Part (d) Now, we consider

$$
\begin{aligned}
& \widehat{\Delta}_{\widehat{u}_0 \Delta_d u_1} \\
= \; & \sum_{j=0}^{K-1} w(j/K)\widehat{\Gamma}_{\widehat{u}_0 \Delta_d u_1}(j) \\
= \; & \sum_{j=0}^{K-1} w(j/K)\left[\widehat{\Gamma}_{\widehat{u}_0 u_1}(j) - \widehat{\Gamma}_{\widehat{u}_0 u_1}(j+d)\right] \\
= \; & \sum_{j=d}^{K+d-1} \left[w(j/K) - w((j-d)/K)\right]\widehat{\Gamma}_{\widehat{u}_0 u_1}(j) + \sum_{j=0}^{d-1} w(j/K)\widehat{\Gamma}_{\widehat{u}_0 u_1}(j) \\
& - \sum_{j=K}^{K+d-1} w(j/K)\widehat{\Gamma}_{\widehat{u}_0 u_1}(j) \\
= \; & \sum_{j=d}^{K+d-1} \Delta_d w(j/K)\widehat{\Gamma}_{\widehat{u}_0 u_1}(j) + \sum_{j=1}^{d-1} w(j/K)\widehat{\Gamma}_{u_0 u_1}(j) + O_p(T^{-1}) \\
& + O_p(T^{-1/2}K^{-1}) + O_p(T^{-1/2}K^{-2}). \qquad (7.7)
\end{aligned}
$$

since

$$
\begin{aligned}
\widehat{\Gamma}_{\widehat{u}_0 u_1}(j) &= \widehat{\Gamma}_{u_0 u_1}(j) - \left(\widehat{A} - A\right)\widehat{\Gamma}_{x u_1}(j) \\
&= \widehat{\Gamma}_{u_0 u_1}(j) - \left(\widehat{A}_1 - A_1\right)\widehat{\Gamma}_{x_1 u_1}(j) - \left(\widehat{A}_2 - A_2\right)\widehat{\Gamma}_{x_2 u_1}(j) \\
&= O_p(T^{-1/2}) \qquad\qquad \text{for } K \leq j \leq K+d-1, \\
\widehat{\Gamma}_{\widehat{u}_0 u_1}(j) &= \widehat{\Gamma}_{u_0 u_1}(j) + O_p(T^{-1}) \qquad \text{for } 1 \leq j \leq d-1,
\end{aligned}
$$

$w(j/K) = O(K^{-2})$ for $K \leq j \leq K+d-1$ and $\widehat{\Gamma}_{\widehat{u}_0 u_1}(0) = T^{-1}\widehat{U}_0' U_1 = T^{-1}\widehat{U}_0' X_1 = 0$ by orthogonality condition of OLS residuals and the tricks as in the proof of part (a).

Then, (7.7) can be further rewritten as

$$
\sum_{j=d}^{K+d-1} \Delta_d w(j/K)\widehat{\Gamma}_{u_0 u_1}(j) - \left(\widehat{A} - A\right)\sum_{j=d}^{K+d-1} \Delta_d w(j/K)\widehat{\Gamma}_{x u_1}(j)
+ \sum_{j=0}^{d-1}\widehat{\Gamma}_{u_0 u_1}(j) + O_p(T^{-1/2}K^{-1}) + O_p(T^{-1/2}K^{-2}). \tag{7.8}
$$

The first sum of (7.8) is $O_p(T^{-1/2}K^{-1/2})$ because its mean is zero under Assumption 2.1 (c) and its variance is $O(T^{-1}K^{-1})$. The second term is $O_p(T^{-1}) + O_p(T^{-1/2}K^{-2})$ as shown in the proof of Lemma 7.2 (e). ■

Part (e) We write

$$
\widehat{\Delta}_{\widehat{u}_0 u_2} = \sum_{j=0}^{K-1} w(j/K)\widehat{\Gamma}_{\widehat{u}_0 u_2}(j)
$$

$$= \sum_{j=0}^{K-1} w(j/K)\widehat{\Gamma}_{u_0u_2}(j) - \left(\widehat{A} - A\right) \sum_{j=0}^{K-1} w(j/K)\widehat{\Gamma}_{xu_2}(j).$$

The first sum is $\sum_{j=0}^{K-1} w(j/K)\widehat{\Gamma}_{u_0u_2}(j) = \Delta_{02} + O_p(T^{-1/2}K^{1/2})$ because its mean is Δ_{02} and variance is $O(T^{-1}K)$. The second sum is $O_p(T^{-1/2})$. The first sum dominates the second sum and hence, $\widehat{\Delta}_{0u_2} := \widehat{\Delta}_{\widehat{u}_0u_2} = \Delta_{02} + O_p(T^{-1/2}K^{1/2})$. ■

LEMMA 7.4. Under Assumptions 2.1, 2.2 and 2.3 (d), the followings hold :

(a) $T^{-1}U_2'S(X_2,d) - \widehat{\Delta}_{u_2u_2} := N_{22T} \to_d \int_0^1 dB_2B_2'$,

(b) $T^{-1}U_0'S(X_2,d) - \widehat{\Delta}_{0u_2} := N_{02T} \to_d \int_0^1 dB_0B_2'$,

(c) $T^{-2}X_2'S(X_2,d) \to_d d^{-1}\int_0^1 B_2B_2'$.

PROOF.

Part (a) We observe that

$$\begin{aligned}
& T^{-1}U_2'S(X_2,d) - \widehat{\Delta}_{u_2u_2} \\
= \;& T^{-1}\sum_{t=1}^{T} u_{2,t}S(x_{2,t}',d) - \sum_{j=0}^{K-1} w(j/K)\widehat{\Gamma}_{u_2u_2}(j) \\
= \;& T^{-1}\sum_{t=1}^{T} u_{2,t}S(x_{2,t-d}',d) + \sum_{j=0}^{d-1} \widehat{\Gamma}_{u_2u_2}(j) - \sum_{j=0}^{d-1} w(j/K)\widehat{\Gamma}_{u_2u_2}(j) \\
& - \sum_{j=d}^{K-1} w(j/K)\widehat{\Gamma}_{u_2u_2}(j) \\
= \;& T^{-1}\sum_{t=1}^{T} u_{2,t}S(x_{2,t-d}',d) - \sum_{j=0}^{d-1} \left[w(j/K) - w(0)\right]\widehat{\Gamma}_{u_2u_2}(j)
\end{aligned}$$

$$
\begin{aligned}
& - \sum_{j=d}^{K-1} w(j/K)\widehat{\Gamma}_{u_2u_2}(j) \\
= \;& T^{-1}\sum_{t=1}^{T} u_{2,t}S(x'_{2,t-d}, d) - \sum_{j=d}^{K-1} w(j/K)\widehat{\Gamma}_{u_2u_2}(j) + O_p(T^{-1/2}K^{-1}) \\
= \;& \int_0^1 dB_2B_2' + O_p(T^{-1/2}) + O_p(T^{-1/2}K^{-1}) + o_p(1). \qquad (7.9)
\end{aligned}
$$

The last line is obtained by using Lemma 2.1 (a) and (2.9). Thus, we deduce that

$$
T^{-1}U_2'S(X_2, d) - \widehat{\Delta}_{u_2u_2} \rightarrow_d \int_0^1 dB_2B_2'.
$$

This is the required result of part (a). ■

Part (b) In addition to Lemma 7.3 (e), the result can be followed in the same way as in part (a), so we will not repeat the proof here.

Part (c) We note that

$$
\begin{aligned}
& T^{-2}X_2'S(X_2, d) \\
= \;& d^{-2}\sum_{j=1}^{T}\int_{d(j-1)/T}^{dj/T} \left((T/d)^{-1/2}x_{2,[Tr]}\right)\left((T/d)^{-1/2}S(x'_{2,[Tr]}, d)\right) dr \\
= \;& d^{-2}\int_0^1 \left((T/d)^{-1/2}x_{2,[Tr]}\right)\left((T/d)^{-1/2}S(x'_{2,[Tr]}, d)\right) dr \\
\rightarrow_d \;& d^{-1}\int_0^1 B_2B_2'.
\end{aligned}
$$

The last line is obtained by (2.8), (2.9) and the continuous mapping theorem. ■

LEMMA 7.5. Under Assumptions 2.1, 2.2 and 2.3 (a), the followings hold :

(a) $\widehat{\Omega}_{0b}\widehat{\Omega}_{bb}^{-1}\left[T^{-1}U_b'S(X_b, d) - \widehat{\Delta}_{bb}\right] = \left[O_p(K^{-2}) + O_p(T^{-1/2}K^{-1/2})\right.$

$$+O_p(T^{-1}K^{1/2}) \vdots \Omega_{02}\Omega_{22}^{-1}N_{22T} + O_p(T^{-1/2}) + O_p(T^{-1}K^{3/2}) + o_p(1)\Big],$$

(b) $T^{1/2}\widehat{\Omega}_{0b}\widehat{\Omega}_{bb}^{-1}\left[T^{-1}U_b'S(X_1,d) - \widehat{\Delta}_{b\Delta_d u_1}\right] = O_p(T^{1/2}K^{-2}) + O_p(K^{-1/2})$

$+O_p(T^{-1/2}K^{1/2})$,

(c) $T^{1/2}\left[T^{-1}U_0'S(X_1,d) - \widehat{\Delta}_{0\Delta_d u_1}\right] = T^{-1/2}U_0'U_1 + O_p(K^{-1/2}) \to_d N(0,\Omega_{\varphi\varphi})$.

PROOF.

Part (a) By Lemma 7.2 (f), Lemma 7.3 (a), (b), (c) and Lemma 7.4 (a), we deduce that

$$\begin{aligned}
& \widehat{\Omega}_{0b}\widehat{\Omega}_{bb}^{-1}\left[T^{-1}U_b'S(X_b,d) - \widehat{\Delta}_{bb}\right] \\
= & \left[-d^{-1}\left(\phi_{01} - \Omega_{02}\Omega_{22}^{-1}\phi_{21}\right)\Omega_{11}^{-1} + O_p(T^{-1/2}K^{3/2}) \vdots \Omega_{02}\Omega_{22}^{-1} + o_p(1)\right] \\
& \times \begin{bmatrix} T^{-1}\Delta_d U_1'S(U_1,d) - \widehat{\Delta}_{\Delta_d u_1 \Delta_d u_1} & T^{-1}\Delta_d U_1'S(X_2,d) - \widehat{\Delta}_{\Delta_d u_1 u_2} \\ T^{-1}U_2'S(U_1,d) - \widehat{\Delta}_{u_2 \Delta_d u_1} & T^{-1}U_2'S(X_2,d) - \widehat{\Delta}_{u_2 u_2} \end{bmatrix} \\
= & \left[-d^{-1}\left(\phi_{01} - \Omega_{02}\Omega_{22}^{-1}\phi_{21}\right)\Omega_{11}^{-1} + O_p(T^{-1/2}K^{3/2}) \vdots \Omega_{02}\Omega_{22}^{-1} + o_p(1)\right] \\
& \times \begin{bmatrix} O_p(K^{-2}) + O_p(T^{-1/2}K^{-1}) & O_p(K^{-2}) + O_p(T^{-1/2}) \\ O_p(K^{-2}) + O_p(T^{-1/2}K^{-1/2}) & N_{22T} \end{bmatrix} \\
= & \Big[O_p(K^{-2}) + O_p(T^{-1/2}K^{-1/2}) + O_p(T^{-1}K^{1/2}) \vdots \\
& \Omega_{02}\Omega_{22}^{-1}N_{22T} + O_p(T^{-1/2}) + O_p(T^{-1}K^{3/2}) + o_p(1)\Big].
\end{aligned}$$

This is the required result of part (a) under Assumption 2.3 (a). ■

Part (b) To prove this part, we can obtain the first submatrix of the result in part (a). Thus,

$$\begin{aligned}
& T^{1/2}\widehat{\Omega}_{0b}\widehat{\Omega}_{bb}^{-1}\left[T^{-1}U_b'S(X_1,d)-\widehat{\Delta}_{b\Delta_d u_1}\right] \\
= \; & T^{1/2}\left[O_p(K^{-2})+O_p(T^{-1/2}K^{-1/2})+O_p(T^{-1}K^{1/2})\right] \\
= \; & \left[O_p(T^{1/2}K^{-2})+O_p(K^{-1/2})+O_p(T^{-1/2}K^{1/2})\right].
\end{aligned}$$

■

Part (c) By Lemma 7.3 (d), we find that

$$\begin{aligned}
& T^{1/2}\left[T^{-1}U_0'S(X_1,d)-\widehat{\Delta}_{0\Delta_d u_1}\right] \\
= \; & T^{-1/2}U_0'S(U_1,d)-T^{1/2}\widehat{\Delta}_{0\Delta_d u_1} \\
= \; & T^{-1/2}U_0'S(U_1,d)-T^{1/2}\left[\sum_{j=1}^{d-1}\widehat{\Gamma}_{u_0u_1}(j)+O_p(T^{-1})+O_p(T^{-1/2}K^{-1/2})\right] \\
= \; & T^{-1/2}U_0'U_1+O_p(T^{-1/2})+O_p(K^{-1/2}) \\
\to_d \; & N\left(0,\Omega_{\varphi\varphi}\right).
\end{aligned}$$

The last line is obtained by the Central Limit Theorem. ■

LEMMA 7.6. Under Assumptions 2.1, 2.2 and 2.3 (d), the followings hold :

(a) $U_2'S(P,d)\left(T^{-1/2}\delta_T^{-1}\right) := N_{2pT} \to_d d\int_0^1 dB_2p'$,

(b) $U_0'S(P,d)\left(T^{-1/2}\delta_T^{-1}\right) := N_{0pT} \to_d d\int_0^1 dB_0p'$,

(c) $W_{2,T}^{-1}Z_2'S(Z_2',d)W_{2,T}^{-1} \to_d \begin{bmatrix} d^{-1}\int_0^1 B_2B_2' & \int_0^1 B_2p' \\ \int_0^1 pB_2' & d\int_0^1 pp' \end{bmatrix}.$

PROOF :

Part (a)

$$
\begin{aligned}
U_2' S(P,d)\left(T^{-1/2}\delta_T^{-1}\right) &= d^{-1/2}(T/d)^{-1/2}\sum_{t=1}^{T} u_{2,t} S(p_t', d)\delta_T^{-1} \\
&= d^{-1/2}(T/d)^{-1/2}\sum_{t=1}^{T} u_{2,t} S(p_t', d)\delta_T^{-1} \\
&= d^{-1/2}\sum_{j=1}^{T}\int_{d(j-1)/T}^{dj/T} d\left((T/d)^{-1/2} u_{2,[Tr]}\right) S\left(p_{[Tr]}', d\right)\delta_T^{-1} \\
&= d^{-1/2}\int_0^1 d\left((T/d)^{-1/2} u_{2,[Tr]}\right) S\left(p_{[Tr]}', d\right)\delta_T^{-1} \\
&\rightarrow_d \; d\int_0^1 dB_2 p',
\end{aligned}
$$

by (2.8), (2.9), (3.15) and the continuous mapping theorem. ■

Part (b) The proof of this part can be followed easily from part (a), so we will not repeat here.

Part (c)

$$
W_{2,T}^{-1} Z_2' S(Z_2', d) W_{2,T}^{-1} = \begin{bmatrix} T^{-2} X_2' S(X_2, d) & T^{-3/2} X_2' S(P,d)\delta_T^{-1} \\ T^{-3/2}\delta_T^{-1} P' S(X_2, d) & T^{-1}\delta_T^{-1} P' S(P,d)\delta_T^{-1} \end{bmatrix}.
$$

We consider the limiting process of $T^{-1}\delta_T^{-1}P'S(P,d)\delta_T^{-1}$, firstly,

$$
\begin{aligned}
T^{-1}\delta_T^{-1} P' S(P,d)\delta_T^{-1} &= T^{-1}\sum_{j=1}^{T}\int_{(j-1)/T}^{j/T} \delta_T^{-1} p_{[Tr]} S\left(p_{[Tr]}', d\right)\delta_T^{-1} dr \\
&= T^{-1}\int_0^1 \delta_T^{-1} p_{[Tr]} S\left(p_{[Tr]}', d\right)\delta_T^{-1} dr \\
&\rightarrow_d \; d\int_0^1 pp',
\end{aligned}
$$

by (3.15) and the continuous mapping theorem.

Next, we consider the limiting process of $T^{-3/2}X_2'S(P,d)\delta_T^{-1}$.

$$\begin{aligned} T^{-3/2}X_2'S(P,d)\delta_T^{-1} &= T^{-3/2}\sum_{t=1}^{T} x_{2,t}S(p_t',d)\delta_T^{-1} \\ &= d^{-3/2}\sum_{j=1}^{T}\int_{d(j-1)/T}^{dj/T}(T/d)^{-1/2}x_{2,[Tr]}S\left(p_{[Tr]}',d\right)\delta_T^{-1}dr \\ &= d^{-3/2}\int_0^1 (T/d)^{-1/2}x_{2,[Tr]}S\left(p_{[Tr]}',d\right)\delta_T^{-1}dr \\ &\to_d \int_0^1 B_2P' \end{aligned}$$

by (2.8), (2.9), (3.15) and the continuous mapping theorem. Similarly, we can also show that $T^{-3/2}\delta_T^{-1}P'S(X_2,d)\to_d \int_0^1 pB_2'$. Combining the results, we have

$$W_{2,T}^{-1}Z_2'S(Z_2',d)W_{2,T}^{-1}\to_d \begin{bmatrix} d^{-1}\int_0^1 B_2B_2' & \int_0^1 B_2p' \\ \int_0^1 pB_2' & d\int_0^1 pp' \end{bmatrix}. \quad \blacksquare$$

PROOF OF THEOREM 3.1. Recall that $\widehat{A}^+ - A = \left(U_0^{+\prime}S(X,d) - T\widehat{\Delta}_{0x}^+\right)$ $\times\left(X'S(X,d)\right)^{-1}$ where $U_0^{+\prime} = U_0' - \widehat{\Omega}_{0x}\widehat{\Omega}_{xx}^{-1}\Delta_d X'$ and $\widehat{\Delta}_{0x}^+ = \widehat{\Delta}_{0x} - \widehat{\Omega}_{0x}\widehat{\Omega}_{xx}^{-1}\widehat{\Delta}_{xx}$.

$$\begin{aligned} \left(\widehat{A}^+ - A\right)D &= \left[\left(\widehat{A}^+ - A\right)D_1 \,\vdots\, \left(\widehat{A}^+ - A\right)D_2\right] \\ &= \left(U_0^{+\prime}S(X,d) - T\widehat{\Delta}_{0x}^+\right)D\left(D'X'S(X,d)D\right)^{-1}D'D. \end{aligned}$$

By partitioned matrix inversion, we have

$$
\begin{aligned}
(D'X'S(X,d)D)\,D'D_1 &= \begin{bmatrix} X_1'S(X_1,d) & X_1'S(X_2,d) \\ X_2'S(X_1,d) & X_2'S(X_2,d) \end{bmatrix}^{-1} \begin{bmatrix} I \\ 0 \end{bmatrix} \\
&= \begin{bmatrix} (X_1'Q_2S(X_1,d))^{-1} \\ -(X_2'S(X_2,d))^{-1}(X_2'S(X_1,d))(X_1'Q_2S(X_1,d))^{-1} \end{bmatrix},
\end{aligned}
$$

$$
\begin{aligned}
(D'X'S(X,d)D)\,D'D_2 &= \begin{bmatrix} X_1'S(X_1,d) & X_1'S(X_2,d) \\ X_2'S(X_1,d) & X_2'S(X_2,d) \end{bmatrix}^{-1} \begin{bmatrix} 0 \\ I \end{bmatrix} \\
&= \begin{bmatrix} -(X_1'S(X_1,d))^{-1}(X_1'S(X_2,d))(X_2'Q_1S(X_2,d))^{-1} \\ (X_2'Q_1S(X_2,d))^{-1} \end{bmatrix},
\end{aligned}
$$

where $Q_i = I - S(X_i,d)\,(X_i'S(X_i,d))^{-1}\,X_i'$ for $i = 1, 2$.

$$
\begin{aligned}
& T^{1/2}\left(\widehat{A}^{+} - A\right)D_1 \\
= \;& T^{1/2}\left[T^{-1}U_0'S(X,d) - \widehat{\Omega}_{0x}\widehat{\Omega}_{xx}^{-1}\left(T^{-1}\Delta_d X'S(X,d)\right) - \left(\widehat{\Delta}_{0x} - \widehat{\Omega}_{0x}\widehat{\Omega}_{xx}^{-1}\widehat{\Delta}_{xx}\right)\right] \\
& \times D\begin{bmatrix} (T^{-1}X_1'Q_2S(X_1,d))^{-1} \\ -T^{-1}(T^{-2}X_2'S(X_2,d))^{-1}(T^{-1}X_2'S(X_1,d))(T^{-1}X_1'Q_2S(X_1,d))^{-1} \end{bmatrix} \\
= \;& T^{1/2}\left[T^{-1}U_0'S(X,d) - \widehat{\Omega}_{0x}D\left(D'\widehat{\Omega}_{xx}D\right)^{-1}D'\left(T^{-1}\Delta_d X'S(X,d)\right)\right. \\
& \left. - \left(\widehat{\Delta}_{0x} - \widehat{\Omega}_{0x}\widehat{\Omega}_{xx}^{-1}\widehat{\Delta}_{xx}\right)\right]\left[D_1\left(T^{-1}X_1'S(X_1,d)\right)^{-1} + D_2O_p(T^{-1})\right] \\
= \;& T^{1/2}\left[T^{-1}U_0'S(X_1,d) - \widehat{\Omega}_{0b}\widehat{\Omega}_{bb}^{-1}D'\left(T^{-1}\Delta_d X'S(X_1,d)\right)\right.
\end{aligned}
$$

$$
\begin{aligned}
& -\left(\widehat{\Delta}_{0\Delta_d u_1} - \widehat{\Omega}_{0b}\widehat{\Omega}_{bb}^{-1}\widehat{\Delta}_{b\Delta_d u_1}\right)\Big] \left(T^{-1}X_1'S(X_1,d)\right)^{-1} \\
& +\Big[T^{-1}U_0'S(X_2,d) - \widehat{\Omega}_{0b}\widehat{\Omega}_{bb}^{-1}D'\left(T^{-1}\Delta_d X'S(X_2,d)\right) \\
& -\left(\widehat{\Delta}_{0u_2} - \widehat{\Omega}_{0b}\widehat{\Omega}_{bb}^{-1}\widehat{\Delta}_{bu_2}\right)\Big] O_p(T^{-1/2}) \\
= \; & T^{1/2}\left[\left(T^{-1}U_0'S(X_1,d) - \widehat{\Delta}_{0\Delta_d u_1}\right) - \widehat{\Omega}_{0b}\widehat{\Omega}_{bb}^{-1}\left(T^{-1}U_b'S(X_1,d) - \widehat{\Delta}_{b\Delta_d u_1}\right)\right] \\
& \times\left(T^{-1}X_1'S(X_1,d)\right)^{-1} \\
& +\left[\left(T^{-1}U_0'S(X_2,d) - \widehat{\Delta}_{0u_2}\right) - \widehat{\Omega}_{0b}\widehat{\Omega}_{bb}^{-1}\left(T^{-1}U_b'S(X_2,d) - \widehat{\Delta}_{bu_2}\right)\right] O_p(T^{-1/2}) \\
= \; & \Big[T^{1/2}\left(T^{-1}U_0'S(X_1,d) - \widehat{\Delta}_{0\Delta_d u_1}\right) + O_p(T^{1/2}K^{-2}) + O_p(K^{-1/2}) \\
& + O_p(T^{-1/2}K^{1/2})\Big]\left(T^{-1}X_1'S(X_1,d)\right)^{-1} \\
& +\left[O_p(T^{-1/2}) + O_P(T^{-1}) + O_p(T^{-3/2}K^{3/2})\right].
\end{aligned}
$$

The last line is obtained by Lemma 7.4 (b), Lemma 7.5 (a) and (b) and thus,

$$
\begin{aligned}
& T^{1/2}\left(\widehat{A}^{+} - A\right)D_1 \\
= \; & T^{1/2}\left(T^{-1}U_0'S(U_1,d) - \widehat{\Delta}_{0\Delta_d u_1}\right)\left(T^{-1}X_1'S(X_1,d)\right)^{-1} + O_p(T^{1/2}K^{-2}) \\
& + O_p(K^{-1/2}) + O_p(T^{-1/2}K^{1/2}) + O_p(T^{-1/2}) + O_p(T^{-1}) + O_p(T^{-3/2}K^{3/2}) \\
\to_d \; & N\left(0, \left(I \otimes \left(\sum_{j=0}^{d-1}\Gamma_{11}(j)\right)^{-1}\right)\Omega_{\varphi\varphi}\left(I \otimes \left(\sum_{j=0}^{d-1}\Gamma_{11}'(j)\right)^{-1}\right)\right),
\end{aligned}
$$

by Lemma 7.5 (c) when Assumption 2.3 (b) is satisfied.

Next,

$$
T\left(\widehat{A}^{+} - A\right)D_2
$$

$$
\begin{aligned}
&= \Big[T^{-1}U_0'S(X,d) - \widehat{\Omega}_{0x}\widehat{\Omega}_{xx}^{-1}\left(T^{-1}\Delta_d X'S(X,d)\right) - \left(\widehat{\Delta}_{0x} - \widehat{\Omega}_{0x}\widehat{\Omega}_{xx}^{-1}\widehat{\Delta}_{xx}\right)\Big] \\
&\quad \times D \begin{bmatrix} -\left(T^{-1}X_1'S(X_1,d)\right)^{-1}\left(T^{-1}X_1'S(X_2,d)\right)\left(T^{-2}X_2'Q_1S(X_2,d)\right)^{-1} \\ \left(T^{-2}X_2'Q_1S(X_2,d)\right)^{-1} \end{bmatrix} \\
&= \Big[T^{-1}U_0'S(X,d) - \widehat{\Omega}_{0x}D\left(D'\widehat{\Omega}_{xx}D\right)^{-1}D'\left(T^{-1}\Delta_d X'S(X,d)\right) \\
&\quad - \left(\widehat{\Delta}_{0x} - \widehat{\Omega}_{0x}\widehat{\Omega}_{xx}^{-1}\widehat{\Delta}_{xx}\right)\Big] \\
&\quad \times \Big[-D_1\left(T^{-1}X_1'S(X_1,d)\right)^{-1}\left(T^{-1}X_1'S(X_2,d)\right)\left(T^{-2}X_2'Q_1S(X_2,d)\right)^{-1} \\
&\quad + D_2\left(T^{-2}X_2'Q_1S(X_2,d)\right)^{-1}\Big] \\
&= -\Big[\left(T^{-1}U_0'S(U_1,d) - \widehat{\Delta}_{0\Delta_d u_1}\right) - \widehat{\Omega}_{0b}\widehat{\Omega}_{bb}^{-1}\left(T^{-1}U_b'S(U_1,d) - \widehat{\Delta}_{b\Delta_d u_1}\right)\Big] \\
&\quad \times \left(T^{-1}X_1'S(X_1,d)\right)^{-1}\left(T^{-1}X_1'S(X_2,d)\right)\left(T^{-2}X_2'Q_1S(X_2,d)\right)^{-1} \\
&\quad + \Big[\left(T^{-1}U_0'S(X_2,d) - \widehat{\Delta}_{0u_2}\right) - \widehat{\Omega}_{0b}\widehat{\Omega}_{bb}^{-1}\left(T^{-1}U_b'S(X_2,d) - \widehat{\Delta}_{bu_2}\right)\Big] \\
&\quad \times \left(T^{-2}X_2'Q_1S(X_2,d)\right)^{-1} \\
&= \Big[O_p(T^{-1/2}) + O_p(T^{-1/2}K^{-1/2}) + O_p(T^{-1}K^{1/2}) + O_p(K^{-2})\Big]O_p(1) \\
&\quad + \Big[(N_{02T} + o_p(1)) - \Omega_{02}\Omega_{22}^{-1}(N_{22T} + o_p(1)) + O_p(T^{-1/2}) + O_p(T^{-1}K^{3/2})\Big] \\
&\quad \times \left(T^{-2}X_2'S(X_2,d)\right)^{-1} \\
&\to_d d\left(\int_0^1 dB_0B_2' - \Omega_{02}\Omega_{22}^{-1}\int_0^1 dB_2B_2'\right)\left(\int_0^1 B_2B_2'\right)^{-1} \\
&= d\left(\int_0^1 dB_{0.2}B_2'\right)\left(\int_0^1 B_2B_2'\right)^{-1},
\end{aligned}
$$

as required for part (b). Note that the results at last three lines are obtained by using Lemma 7.4 and Lemma 7.5 (a), (c). This part of the results holds when

$k \in (0, 2/3)$. Both part (a) and (b) hold when Assumption 2.3 (a) is satisfied. ■

PROOF OF COROLLARY 3.2. The proof of this corollary can be followed directly from that of Theorem 3.1 and we can see that

$$
\begin{aligned}
& T\left(\widehat{A}_2^{+} - A_2\right) \\
= \; & \left[\left(T^{-1}U_0'S(X_2,d) - \widehat{\Delta}_{0u_2}\right) - \widehat{\Omega}_{0b}\widehat{\Omega}_{bb}^{-1}\left(T^{-1}U_b'S(X_2,d) - \widehat{\Delta}_{bu_2}\right)\right] \\
& \times \left(T^{-2}X_2'S(X_2,d)\right)^{-1} \\
= \; & \left[\left(T^{-1}U_0'S(X_2,d) - \widehat{\Delta}_{0u_2}\right) - \widehat{\Omega}_{02}\widehat{\Omega}_{22}^{-1}\left(T^{-1}U_2'S(X_2,d) - \widehat{\Delta}_{u_2u_2}\right)\right] \\
& \times \left(T^{-2}X_2'S(X_2,d)\right)^{-1} \\
= \; & \left[(N_{02T} + o_p(1)) - \Omega_{02}\Omega_{22}^{-1}(N_{22T} + o_p(1)) + O_p(T^{-1/2}) + O_p(T^{-1}K^{3/2})\right] \\
& \times \left(T^{-2}X_2'S(X_2,d)\right)^{-1} \\
\to_d \; & d\left(\int_0^1 dB_0B_2' - \Omega_{02}\Omega_{22}^{-1}\int_0^1 dB_2B_2'\right)\left(\int_0^1 B_2B_2'\right)^{-1} \\
= \; & d\left(\int_0^1 dB_{0.2}B_2'\right)\left(\int_0^1 B_2B_2'\right)^{-1}.
\end{aligned}
$$

Note that this corollary holds for a wider bandwidth, that is, Assumption 2.3(d), because the absence of stationary component induces the removal of the term $O_p(T^{-1}K^{3/2})$ and also inclusion of stationary component in the model captures the bandwidth expansion rate of $[1/4, 2/3)$. Hence, the exclusion of stationary component can allow the FM-SEA estimator for a wider bandwidth.

■

PROOF OF COROLLARY 3.3. This proof can be followed directly from the proof of Theorem 3.1 and the absence of submatrix for non-stationary component can show the result directly. Unlike the case of the exclusion of non-stationary component in the previous corollary, the bandwidth cannot be expanded to $k \in (0,1)$. This is because the term $O_p(T^{1/2}K^{-2})$ cannot be removed even if the non-stationary component is absent. Furthermore, a too low bandwidth expansion rate does not guarantee the convergence of FM-SEA estimates to OLS estimates.

PROOF OF COROLLARY 3.4. The proof can be followed directly from the proof of Theorem 3.1, so we will not repeat the proof here.

PROOF OF THEOREM 3.5. Recall that $\left(\widehat{\Phi}^{+}-\Phi\right) = \left(U_0^{+\prime}S(Z,d) - \left[T\widehat{\Delta}_{0x}^{+}\vdots 0\right]\right)(Z'S(Z,d))^{-1}$ where $U_0^{+\prime} = U_0' - \widehat{\Omega}_{0x}\widehat{\Omega}_{xx}^{-1}\Delta_d X'$ and $\widehat{\Delta}_{0x}^{+} = \widehat{\Delta}_{0x} - \widehat{\Omega}_{0x}\widehat{\Omega}_{xx}^{-1}\widehat{\Delta}_{xx}$. We define $D_L = \begin{bmatrix} D_1 \\ 0 \end{bmatrix}$ and $D_M = \begin{bmatrix} D_2 & F \\ 0 & I \end{bmatrix}$.

Then, $D = \begin{bmatrix} D_L, & D_M \end{bmatrix}$.

$$\begin{aligned}
\left(\widehat{\Phi}^{+}-\Phi\right) DW_T &= \left[\left(\widehat{\Phi}^{+}-\Phi\right) D_L \vdots \left(\widehat{\Phi}^{+}-\Phi\right) D_M\right] W_T \\
&= \left(U_0^{+\prime}S(Z,d) - \left[T\widehat{\Delta}_{0x}^{+}\vdots 0\right]\right) D \left(D'Z'S(Z,d)D\right)^{-1} D'DW_T.
\end{aligned}$$

By partitioned matrix inversion, we have

$$
\begin{aligned}
\left(D'Z'S(Z,d)D\right)D'D_L &= \begin{bmatrix} Z_1'S(Z_1,d) & Z_1'S(Z_2,d) \\ Z_2'S(Z_1,d) & Z_2'S(Z_2,d) \end{bmatrix}^{-1} \begin{bmatrix} I \\ 0 \end{bmatrix} \\
&= \begin{bmatrix} \left(Z_1'Q_2S(Z_1,d)\right)^{-1} \\ -\left(Z_2'S(Z_2,d)\right)^{-1}\left(Z_2'S(Z_1,d)\right)\left(Z_1'Q_2S(Z_1,d)\right)^{-1} \end{bmatrix},
\end{aligned}
$$

$$
\begin{aligned}
\left(D'Z'S(Z,d)D\right)D'D_M &= \begin{bmatrix} Z_1'S(Z_1,d) & Z_1'S(Z_2,d) \\ Z_2'S(Z_1,d) & Z_2'S(Z_2,d) \end{bmatrix}^{-1} \begin{bmatrix} 0 \\ I \end{bmatrix} \\
&= \begin{bmatrix} -\left(Z_1'S(Z_1,d)\right)^{-1}\left(Z_1'S(Z_2,d)\right)\left(Z_2'Q_1S(Z_2,d)\right)^{-1} \\ \left(Z_2'Q_1S(Z_2,d)\right)^{-1} \end{bmatrix},
\end{aligned}
$$

where $Q_i = I - S(Z_i,d)\left(Z_i'S(Z_i,d)\right)^{-1}Z_i'$ for $i = 1,2$.

Now,

$$
\begin{aligned}
&\left(\widehat{\Phi}^{+} - \Phi\right)D_L W_{1,T} \\
= \;& T^{1/2}\left(\widehat{A}^{+} - A\right)D_1 \\
= \;& T^{1/2}\left[T^{-1}U_0'S(Z,d) - \widehat{\Omega}_{0x}\widehat{\Omega}_{xx}^{-1}\left(T^{-1}\Delta_d X'S(Z,d)\right) - \left(\left(\widehat{\Delta}_{0x} - \widehat{\Omega}_{0x}\widehat{\Omega}_{xx}^{-1}\widehat{\Delta}_{xx}\right)\vdots 0\right)\right]D \\
& \times \begin{bmatrix} \left(T^{-1}Z_1'Q_2S(Z_1,d)\right)^{-1} \\ -W_{2,T}^{-1}\left(W_{2,T}^{-1}Z_2'S(Z_2,d)W_{2,T}^{-1}\right)^{-1}\left(W_{2,T}^{-1}Z_2'S(Z_1,d)\right)\left(T^{-1}Z_1'Q_2S(Z_1,d)\right)^{-1} \end{bmatrix}
\end{aligned}
$$

$$
\begin{aligned}
&= T^{1/2}\Big[T^{-1}U_0'S(Z,d) - \widehat{\Omega}_{0x}\,[D_1,D_2]\left([D_1,D_2]'\,\widehat{\Omega}_{xx}\,[D_1,D_2]\right)^{-1}[D_1,D_2]' \\
&\quad \times \left(T^{-1}\Delta_d X'S(Z,d)\right) - \left(\left(\widehat{\Delta}_{0x} - \widehat{\Omega}_{0x}\widehat{\Omega}_{xx}^{-1}\widehat{\Delta}_{xx}\right) \vdots\, 0\right)\Big]\, D_L\left(T^{-1}Z_1'S(Z_1,d)\right)^{-1} \\
&\quad + \Big[T^{-1}U_0'S(Z,d) - \widehat{\Omega}_{0x}\,[D_1,D_2]\left([D_1,D_2]'\,\widehat{\Omega}_{xx}\,[D_1,D_2]\right)^{-1}[D_1,D_2]' \\
&\quad \times \left(T^{-1}\Delta_d X'S(Z,d)\right) - \left(\left(\widehat{\Delta}_{0x} - \widehat{\Omega}_{0x}\widehat{\Omega}_{xx}^{-1}\widehat{\Delta}_{xx}\right) \vdots\, 0\right)\Big]\, D_M W_{2,T}^{-1} O_p(1) \\
&= T^{1/2}\left[\left(T^{-1}U_0'S(X_1,d) - \widehat{\Delta}_{0\Delta_d u_1}\right) - \widehat{\Omega}_{0b}\widehat{\Omega}_{bb}^{-1}\left(T^{-1}U_b'S(X_1,d) - \widehat{\Delta}_{b\Delta_d u_1}\right)\right] \\
&\quad \times \left(T^{-1}Z_1'S(Z_1,d)\right)^{-1} \\
&\quad + \left[\left(T^{-1}U_0'S(X_2,d) - \widehat{\Delta}_{0u_2}\right) - \widehat{\Omega}_{0b}\widehat{\Omega}_{bb}^{-1}\left(T^{-1}U_b'S(X_2,d) - \widehat{\Delta}_{bu_2}\right)\right] O_p(T^{-1/2}) \\
&\quad + \left[U_0'S(P,d)\left(T^{-1/2}\delta_T^{-1}\right) - \widehat{\Omega}_{0b}\widehat{\Omega}_{bb}^{-1}U_b'S(P,d)\left(T^{-1/2}\delta_T^{-1}\right)\right] O_p(T^{-1/2}) \\
&= \Big[T^{1/2}\left(T^{-1}U_0'S(X_1,d) - \widehat{\Delta}_{0\Delta_d u_1}\right) + O_p(T^{1/2}K^{-2}) + O_p(K^{-1/2}) \\
&\quad + O_p(T^{-1/2}K^{1/2})\Big]\left(T^{-1}X_1'S(X_1,d)\right)^{-1} \\
&\quad + \left[O_p(T^{-1/2}) + O_p(T^{-1}) + O_p(T^{-3/2}K^{3/2}) + O_p(T^{-1/2})\right].
\end{aligned}
$$

The last line of this proof is obtained by the conditions in the proof of Theorem 3.1 (a) and Lemma 7.6. Since we can follow the proof of Theorem 3.1 (a) directly, we have

$$
\begin{aligned}
& T^{1/2}\left(\widehat{A}^{+} - A\right) D_1 \\
\to_d\ & N\left(0, \left(I \otimes \left(\sum_{j=0}^{d-1}\Gamma_{11}(j)\right)^{-1}\right)\Omega_{\varphi\varphi}\left(I \otimes \left(\sum_{j=0}^{d-1}\Gamma_{11}'(j)\right)^{-1}\right)\right),
\end{aligned}
$$

when Assumption 2.3 (b) is satisfied.

Next,

$$
\begin{aligned}
& \left(\widehat{\Phi}^{+} - \Phi\right) D_M W_{2,T} \\
= & \left[U_0' S(Z,d) - \widehat{\Omega}_{0x} \widehat{\Omega}_{xx}^{-1} \left(\Delta_d X' S(X,d)\right) - \left(T\left(\widehat{\Delta}_{0x} - \widehat{\Omega}_{0x}\widehat{\Omega}_{xx}^{-1}\widehat{\Delta}_{xx}\right) \vdots 0\right)\right] D \\
& \times \begin{bmatrix} T^{-1} I_{m_1} & 0 \\ 0 & W_{2,T}^{-1} \end{bmatrix} \\
& \times \begin{bmatrix} -\left(T^{-1} Z_1' S(Z_1,d)\right)^{-1} \left(T^{-1} Z_1' S(Z_2,d) W_{2,T}^{-1}\right) \left(W_{2,T}^{-1} Z_2' Q_1 S(Z_2^{-1},d) W_{2,T}^{-1}\right)^{-1} \\ \left(W_{2,T}^{-1} Z_2' Q_1 S(Z_2,d) W_{2,T}^{-1}\right)^{-1} \end{bmatrix} \\
= & \left[T^{-1} U_0' S(Z,d) - \widehat{\Omega}_{0x} \left[D_1, D_2\right] \left(\left[D_1, D_2\right]' \widehat{\Omega}_{xx} \left[D_1, D_2\right]\right)^{-1} \left[D_1, D_2\right]'\right. \\
& \left.\times \left(T^{-1} \Delta_d X' S(X,d)\right) - \left(\widehat{\Delta}_{0x} - \widehat{\Omega}_{0x}\widehat{\Omega}_{xx}^{-1}\widehat{\Delta}_{xx}\right)\right] D_1 O_p(1) \\
& + \left[U_0' S(Z,d) - \widehat{\Omega}_{0x} \left[D_1, D_2\right] \left(\left[D_1, D_2\right]' \widehat{\Omega}_{xx} \left[D_1, D_2\right]\right)^{-1} \left[D_1, D_2\right]' \Delta_d X' S(X,d)\right. \\
& \left. - \left(T\left(\widehat{\Delta}_{0x} - \widehat{\Omega}_{0x}\widehat{\Omega}_{xx}^{-1}\widehat{\Delta}_{xx}\right) \vdots 0\right)\right] D_M W_{2,T}^{-1} \left(W_{2,T}^{-1} Z_2' Q_1 S(Z_2,d) W_{2,T}^{-1}\right)^{-1} \\
= & \left[\left(T^{-1} U_0' S(U_1,d) - \widehat{\Delta}_{0\Delta_d u_1}\right) - \widehat{\Omega}_{0b}\widehat{\Omega}_{bb}^{-1} \left(T^{-1} U_b' S(U_1,d) - \widehat{\Delta}_{b\Delta_d u_1}\right)\right] O_p(1) \\
& + \left[\left(T^{-1} U_0' S(X_2,d) - \widehat{\Delta}_{0u_2}\right) - \widehat{\Omega}_{0b}\widehat{\Omega}_{bb}^{-1} \left(T^{-1} U_b' S(X_2,d) - \widehat{\Delta}_{bu_2}\right) \vdots \right. \\
& \left. U_0' S(P,d) \left(T^{-1/2}\delta_T^{-1}\right) - \widehat{\Omega}_{0b}\widehat{\Omega}_{bb}^{-1} U_b' S(P,d) \left(T^{-1/2}\delta_T^{-1}\right)\right] \\
& \times \left(W_{2,T}^{-1} Z_2' Q_1 S(Z_2,d) W_{2,T}^{-1}\right)^{-1} \\
= & \left[O_p(T^{-1/2}) + O_p(T^{-1/2}K^{-1/2}) + O_p(T^{-1}K^{1/2}) + O_p(K^{-2})\right] O_p(1) \\
& + \left[\left(N_{02T} + o_p(1)\right) - \Omega_{02}\Omega_{22}^{-1}\left(N_{22T} + o_p(1)\right) + O_p(T^{-1}K^{3/2}) \vdots \right. \\
& \left. \left(N_{0pT} + o_p(1)\right) - \Omega_{02}\Omega_{22}^{-1}\left(N_{2pT} + o_p(1)\right)\right] \left(W_{2,T}^{-1} Z_2' Q_1 S(Z_2,d) W_{2,T}^{-1}\right)^{-1}
\end{aligned}
$$

$$
\begin{aligned}
\to_d \quad & \left[\int_0^1 dB_0 B_2' - \Omega_{02}\Omega_{22}^{-1}\int_0^1 dB_2 B_2' \vdots d\int_0^1 dB_0 p' - d\Omega_{02}\Omega_{22}^{-1}\int_0^1 dB_2 p'\right] \\
& \times \begin{bmatrix} d^{-1}\int_0^1 B_2 B_2' & \int_0^1 B_2 p' \\ \int_0^1 p B_2' & d\int_0^1 pp' \end{bmatrix}^{-1} \\
= \quad & \left[\int_0^1 dB_{0.2}B_2' \vdots d\int_0^1 dB_{0.2}p'\right] \begin{bmatrix} d^{-1}\int_0^1 B_2 B_2' & \int_0^1 B_2 p' \\ \int_0^1 p B_2' & d\int_0^1 pp' \end{bmatrix}^{-1}
\end{aligned}
$$

as required for part (b). The result of last line is obtained by Lemma 7.6. This part of results holds when . Both part (a) and (b) hold when Assumption 2.3 (a) is satisfied. ■

PROOF OF THEOREM 4.1. As $X' = \left[X_1' \vdots X_2'\right]' = \left[Z', P_{1,-1}' \vdots P_{2,-1}'\right]$, we can obtain the result of $\widehat{A}^+ - A$ by (4.12), that is, $\widehat{A}^+ - A = \left(V_0^{+\prime} S(X', d) - T\widehat{\Delta}_{0x}^+\right) \times \left(X' S(X, d)\right)^{-1}$.

By partitioned matrix inversion of $X'S(X,d)$, we have

$$
\begin{aligned}
& T^{1/2}\left(\widehat{A}_1^+ - A_1\right) \\
= \quad & T^{1/2}\left[\left(T^{-1}V_0' S(X_1, d) - \widehat{\Delta}_{0\Delta_d x_1}\right) - \widehat{\Omega}_{0b}\widehat{\Omega}_{bb}^{-1}\left(T^{-1}V_b' S(X_1, d) - \widehat{\Delta}_{b\Delta_d x_1}\right)\right] \\
& \times \left(T^{-1}X_1' Q_2 S(X_1, d)\right)^{-1} \\
& - T^{1/2}\left[\left(T^{-1}V_0' S(P_{2,-1}, d) - \widehat{\Delta}_{0v_2}\right) - \widehat{\Omega}_{0b}\widehat{\Omega}_{bb}^{-1}\left(T^{-1}V_b' S(P_{2,-1}, d) - \widehat{\Delta}_{bv_2}\right)\right] \\
& \times \left[T^{-1}\left(T^{-2}X_2' S(X_2, d)\right)^{-1}\left(T^{-1}X_2' S(X_1, d)\right)\left(T^{-1}X_1' Q_2 S(X_1, d)\right)^{-1}\right] \\
= \quad & T^{1/2}\left[\left(T^{-1}V_0' S(X_1, d) - \widehat{\Delta}_{0\Delta_d v_1}\right) - \widehat{\Omega}_{0b}\widehat{\Omega}_{bb}^{-1}\left(T^{-1}V_b' S(X_1, d) - \widehat{\Delta}_{b\Delta_d v_1}\right)\right]
\end{aligned}
$$

$$
\begin{aligned}
& \times \left(T^{-1}X_1'(X_1,d)\right)^{-1} + O_p(T^{-1/2}) \\
= \; & \left[T^{-1/2}V_0'X_1 + O_p(K^{-1/2}) + O_p(T^{1/2}K^{-2}) + O_p(T^{-1/2}K^{1/2})\right] \\
& \times \left(T^{-1}X_1'S(X_1,d)\right)^{-1} + O_p(T^{-1/2}).
\end{aligned}
$$

where $Q_j = I - S(X_1,d)\left(X_j'S(X_j,d)\right)^{-1}X_j'$, for $j = 1,2$, $v_{b,t} = \left[\Delta_d v_{1,t}', v_{2,t}'\right]'$.

The last line is obtained by Lemma 7.5 (b) and (c) as in the proof of Theorem 3.1 (a). Thus,

$$
\begin{aligned}
& T^{1/2}\left(\widehat{A}_1^+ - A_1\right) \\
\to_d \; & N\left(0, \left(I \otimes \left(\sum_{j=0}^{d-1}\Gamma_{11}(j)\right)^{-1}\right)(\Sigma_{00}\otimes\Sigma_{11})\left(I \otimes \left(\sum_{j=0}^{d-1}\Gamma_{11}'(j)\right)^{-1}\right)\right) \\
= \; & N\left(0, \left(\Sigma_{00} \otimes \left(\sum_{j=0}^{d-1}\Gamma_{11}(j)\right)^{-1}\Sigma_{11}\left(\sum_{j=0}^{d-1}\Gamma_{11}'(j)\right)^{-1}\right)\right).
\end{aligned}
$$

when Assumption 2.3 (b) is satisfied.

Next, we consider the limiting distribution of non-stationary component.

$$
\begin{aligned}
& T\left(\widehat{A}_2^+ - A_2\right) \\
= \; & -\left[\left(T^{-1}V_0'S(X_1,d) - \widehat{\Delta}_{0\Delta_d x_1}\right) - \widehat{\Omega}_{0b}\widehat{\Omega}_{bb}^{-1}\left(T^{-1}V_b'S(X_1,d) - \widehat{\Delta}_{b\Delta_d x_1}\right)\right] \\
& \times \left(T^{-1}X_1'S(X_1,d)\right)^{-1}\left(T^{-1}X_1'S(X_2,d)\right)\left(T^{-2}X_2'Q_1S(X_2,d)\right)^{-1} \\
& + \left[\left(T^{-1}V_0'S(X_2,d) - \widehat{\Delta}_{0v_2}\right) - \widehat{\Omega}_{0b}\widehat{\Omega}_{bb}^{-1}\left(T^{-1}V_b'S(X_2,d) - \widehat{\Delta}_{bv_2}\right)\right] \\
& \times \left(\left(T^{-2}X_2'Q_1S(X_2,d)\right)^{-1}\right.
\end{aligned}
$$

$$
\begin{aligned}
&= \left[O_p(T^{-1/2}) + O_p(T^{-1/2}K^{-1/2}) + O_p(T^{-1}K^{1/2}) + O_p(K^{-2})\right] O_p(1) \\
&\quad + \left[\left(T^{-1}V_0'S(X_2,d) - \widehat{\Delta}_{0v_2}\right) - \widehat{\Omega}_{0b}\widehat{\Omega}_{bb}^{-1}\left(T^{-1}V_b'S(X_2,d) - \widehat{\Delta}_{bv_2}\right)\right] \\
&\quad \times \left((T^{-2}X_2'S(X_2,d)\right)^{-1}.
\end{aligned}
$$

As in the proof of Theorem 3.1 (b),

$$
T\left(\widehat{A}_2^{+} - A_2\right) \to_d d\left(\int_0^1 dB_{0.2}B_2'\right)\left(\int_0^1 B_2B_2'\right)^{-1},
$$

when Assumption 2.3 (c) is satisfied where $B_{0.2} = B_0 - \Omega_{02}\Omega_{22}^{-1}B_2 \equiv BM(\Omega_{0.22})$ and $\Omega_{0.22} = \Omega_{00} - \Omega_{02}\Omega_{22}^{-1}\Omega_{20}$. ■

PROOF OF COROLLARY 4.2. The proof of part (a) is followed easily from the proof of Theorem 4.1 (a) in which the variable $p_{1,t-1}$ vanishes. Also, the component $x_{1,t}$ contains the lagged values of $\Delta_d y_t$, that is, z_t only. In addition, the FM-SEA regression includes the serial correction term for z_t. It follows that the limiting distribution of $\widehat{A}_1^{+}$ is asymptotically normal with bandwidth parameter $k \in (1/4, 1)$.

To prove part (b), we consider the following expression :

$$
\begin{aligned}
v_{2,t} &= \Delta_d p_{2,t-1} \\
&= \Delta_d p_{t-1}
\end{aligned}
$$

$$= \quad (1-L^d)\begin{bmatrix} m_{1,t-1} \\ m_{2,t-1} \\ \vdots \\ m_{d,t-1} \end{bmatrix} = \begin{bmatrix} F_1(L)\Delta_d y_t \\ F_2(L)\Delta_d y_t \\ \vdots \\ F_d(L)\Delta_d y_t \end{bmatrix} = J(L)v_{0,t}, \text{ say,}$$

where when d is odd, $F_1(L) = 1 + L + L^2 + ... + L^{d-1}$,

$$F_j(L) = \begin{cases} \dfrac{\varphi_{j/2}(L)}{1-\exp(\iota\theta_{j/2})L} & \text{for } j = 2, 4, ..., d-1, \\ \dfrac{\varphi_{(j-1)/2}(L)}{1-\exp(\iota\theta_{(j-1)/2})L} & \text{for } j = 3, 5, ..., d, \end{cases}$$

when d is even, $F_1(L) = 1 + L + L^2 + ... + L^{d-1}$, $F_2(L) = 1 - L + L^2 + ... - L^{d-1}$,

$$F_j(L) = \begin{cases} \dfrac{\varphi_{(j-1)/2}(L)}{1-\exp(\iota\theta_{(j-1)/2})L} & \text{for } j = 3, 5, ..., d-1, \\ \dfrac{\varphi_{(j-2)/2}(L)}{1-\exp(\iota\theta_{(j-2)/2})L} & \text{for } j = 4, 6, ..., d. \end{cases}$$

The last line above can be obtained because no element of y_t is seasonally cointegrated and $\Delta_d y_t$ can be expressed in terms of $v_{0,t}$ by (4.5). Explicitly, (4.5) becomes

$$\Delta_d y_t = H_1^* \Delta_d y_{t-1} + H_2^* \Delta_d y_{t-2} + ... + H_{k-d}^* \Delta_d y_{t-k+d} + v_{0,t}.$$

Then,

$$\Delta_d y_t = (I - H_1^* L - H_2^* L^2 - ... - H_{k-d}^* L^{k-d})^{-1} v_{0,t}.$$

Hence, $v_{2,t} = J(L)v_{0,t}$ where

$$J(L) = \begin{bmatrix} F_1(L)(I - H_1^* L - H_2^* L^2 - \dots - H_{k-d}^* L^{k-d})^{-1} \\ F_2(L)(I - H_1^* L - H_2^* L^2 - \dots - H_{k-d}^* L^{k-d})^{-1} \\ \vdots \\ F_d(L)(I - H_1^* L - H_2^* L^2 - \dots - H_{k-d}^* L^{k-d})^{-1} \end{bmatrix}.$$

Then, $\Omega_{22} = J\Omega_{00}J'$ and $B_2(r) = JB_0(r)$ where $J = J(1)$ and $|J(1)| \neq 0$. Hence,

$$\begin{aligned} B_{0.2} &= B_0 - \Omega_{00}J'\left(J\Omega_{00}J'\right)^{-1} JB_0 \\ &= B_0 - B_0 = 0 \qquad\qquad \text{almost surely.} \end{aligned}$$

The limiting distribution of $\widehat{A}_2^+$ given in Theorem 4.1 (b) becomes

$$T\left(\widehat{A}_2^+ - A_2\right) = T\left(\widehat{\Pi}^+ - \Pi\right) \to_d d\left(\int_0^1 dB_{0.2}B_2'\right)\left(\int_0^1 B_2 B_2'\right)^{-1} = 0 \text{ almost surely.}$$

■

PROOF OF COROLLARY 4.3. The result can be obtained easily from the proof of Theorem 4.1 (a) because the submatrix A_2 disappears in the model (4.10).

PROOF OF COROLLARY 4.4. Since the proof of this corollary is a special case of Theorem 4.1, we will not repeat the proof for convenience.

PROOF OF COROLLARY 4.5. Since the proof of this corollary can be followed directly from that of Theorem 4.1, we will not repeat the proof here.

Table 1 Summary Statistics ($d = 1$, $T = 50$ and $T = 100$)

			$T = 50$					
			Coefficients			t-statistics		
μ			MAD	Bias_{avg}	RMSE_{avg}	MAD	Bias_{avg}	RMSE_{avg}
0.0	a_1	FM	0.0726	-0.0008	1.8925	0.1679	-0.0024	4.3597
		OLS	0.0667	-0.0008	1.7310	1.0902	-0.0166	27.2802
	a_2	FM	0.0404	-0.0003	1.1433	0.1650	-0.0001	4.3101
		OLS	0.0377	-0.0001	1.0581	1.1050	0.0012	27.6483
0.4	a_1	FM	0.0754	0.0001	1.9694	0.1610	0.0000	4.2034
		OLS	0.0724	0.0005	1.8800	1.0897	0.0060	27.2401
	a_2	FM	0.0419	0.0089	1.2016	0.1588	0.0344	4.1941
		OLS	0.0423	0.0206	1.1918	1.1448	0.5519	28.5779
0.6	a_1	FM	0.0772	0.0008	2.0215	0.1539	0.0007	4.0608
		OLS	0.0780	0.0008	2.0286	1.0920	0.0078	27.3723
	a_2	FM	0.0429	0.0132	1.2301	0.1517	0.0482	3.9986
		OLS	0.0473	0.0309	1.3289	1.1963	0.7763	29.7429
1.0	a_1	FM	0.0852	0.0011	2.2473	0.1403	0.0019	3.7289
		OLS	0.0950	0.0016	2.4672	1.0984	0.0136	27.5340
	a_2	FM	0.0470	0.0218	1.3733	0.1385	0.0645	3.7157
		OLS	0.0614	0.0512	1.7080	1.2748	1.0561	31.1600

(Continued)

Table 1 (Continued)

			$T = 100$					
			Coefficients			t-statistics		
μ			MAD	Bias_{avg}	RMSE_{avg}	MAD	Bias_{avg}	RMSE_{avg}
0.0	a_1	FM	0.0448	0.0003	0.8147	0.1030	0.0005	1.8625
		OLS	0.0410	0.0002	0.7431	1.1149	0.0068	19.8344
	a_2	FM	0.0189	-0.0002	0.3780	0.0999	-0.0007	1.8060
		OLS	0.0179	-0.0001	0.3553	1.1433	-0.0094	20.2733
0.4	a_1	FM	0.0463	-0.0005	0.8430	0.0986	-0.0012	1.7892
		OLS	0.0447	-0.0007	0.8109	1.1272	-0.0211	20.0654
	a_2	FM	0.0192	0.0022	0.3864	0.0944	0.0109	1.7082
		OLS	0.0202	0.0099	0.4017	1.1978	0.5863	21.1914
0.6	a_1	FM	0.0482	0.0002	0.8802	0.0867	0.0002	1.5819
		OLS	0.0541	-0.0001	0.9796	1.1474	-0.0014	20.4061
	a_2	FM	0.0198	0.0051	0.4022	0.0823	0.0206	1.5107
		OLS	0.0262	0.0206	0.5176	1.3090	1.0169	22.9357
1.0	a_1	FM	0.0498	-0.0004	0.9104	0.0811	-0.0006	1.4794
		OLS	0.0598	0.0001	1.0817	1.1516	0.0043	20.4312
	a_2	FM	0.0200	0.0057	0.4076	0.0755	0.0218	1.3844
		OLS	0.0295	0.0251	0.5804	1.3474	1.1361	23.4323

Table 2 Summary Statistics ($d = 1$, $T = 200$ and $T = 500$)

			$T = 200$					
			Coefficients			t-statistics		
μ			MAD	Bias_{avg}	RMSE_{avg}	MAD	Bias_{avg}	RMSE_{avg}
0.0	a_1	FM	0.0297	-0.0001	0.3770	0.0684	-0.0003	0.8635
		OLS	0.0274	-0.0002	0.3467	1.1113	-0.0082	13.8932
	a_2	FM	0.0088	-0.0001	0.1252	0.0638	-0.0004	0.8074
		OLS	0.0085	0.0000	0.1202	1.1133	-0.0041	13.9871
0.4	a_1	FM	0.0299	0.0002	0.3796	0.0640	0.0002	0.8112
		OLS	0.0295	0.0002	0.3742	1.1165	0.0021	13.9872
	a_2	FM	0.0090	0.0006	0.1278	0.0602	0.0039	0.7652
		OLS	0.0098	0.0049	0.1384	1.1781	0.5880	14.7401
0.6	a_1	FM	0.0301	0.0001	0.3832	0.0597	0.0002	0.7592
		OLS	0.0318	0.0002	0.4040	1.1109	0.0069	13.9687
	a_2	FM	0.0091	0.0008	0.1301	0.0557	0.0048	0.7073
		OLS	0.0111	0.0074	0.1562	1.2225	0.8102	15.1734
1.0	a_1	FM	0.0313	0.0001	0.3975	0.0511	0.0000	0.6470
		OLS	0.0391	0.0001	0.4941	1.1271	0.0050	14.0543
	a_2	FM	0.0091	0.0016	0.1296	0.0464	0.0077	0.5913
		OLS	0.0144	0.0124	0.2006	1.3182	1.1262	16.0870

(Continued)

Table 2 (Continued)

			$T = 500$					
			Coefficients			t-statistics		
μ			MAD	Bias_{avg}	RMSE_{avg}	MAD	Bias_{avg}	RMSE_{avg}
0.0	a_1	FM	0.0178	-0.0002	0.1422	0.0413	-0.0005	0.3291
		OLS	0.0165	-0.0002	0.1315	1.1237	-0.0142	8.9210
	a_2	FM	0.0035	-0.0001	0.0308	0.0384	-0.0001	0.3052
		OLS	0.0034	0.0000	0.0303	1.1347	-0.0009	8.9668
0.4	a_1	FM	0.0176	-0.0001	0.1406	0.0378	-0.0002	0.3016
		OLS	0.0175	-0.0001	0.1401	1.1035	-0.0037	8.7868
	a_2	FM	0.0035	0.0001	0.0309	0.0354	0.0014	0.2821
		OLS	0.0038	0.0020	0.0340	1.1744	0.6026	9.2895
0.6	a_1	FM	0.0179	0.0002	0.1430	0.0356	0.0004	0.2839
		OLS	0.0191	0.0001	0.1523	1.1133	0.0059	8.8435
	a_2	FM	0.0035	0.0001	0.0310	0.0327	0.0013	0.2610
		OLS	0.0043	0.0029	0.0381	1.2150	0.8234	9.5651
1.0	a_1	FM	0.0181	0.0000	0.1447	0.0295	-0.0001	0.2362
		OLS	0.0232	-0.0001	0.1852	1.1172	-0.0055	8.8602
	a_2	FM	0.0035	0.0002	0.0313	0.0272	0.0017	0.2163
		OLS	0.0056	0.0049	0.0498	1.3224	1.1366	10.2308

Table 3 Summary Statistics ($d = 4$, $T = 50$ and $T = 100$)

			$T = 50$					
			Coefficients			t-statistics		
μ			MAD	Bias_{avg}	RMSE_{avg}	MAD	Bias_{avg}	RMSE_{avg}
0.0	a_1	FM	0.1646	0.0020	4.4256	0.2955	0.0024	8.1157
		OLS	0.0829	0.0009	2.1615	0.9401	0.0110	23.6368
	a_2	FM	0.1294	-0.0009	3.5553	0.2886	0.0001	8.1161
		OLS	0.0623	0.0001	1.6533	0.9350	0.0021	23.5046
0.4	a_1	FM	0.1669	-0.0013	4.6820	0.2753	-0.0009	7.6361
		OLS	0.0905	-0.0005	2.3461	0.9464	-0.0066	23.7383
	a_2	FM	0.1342	0.0275	3.8155	0.2765	0.0655	7.7976
		OLS	0.0882	0.0710	2.2948	1.2176	0.9800	29.8502
0.6	a_1	FM	0.1681	0.0005	4.5235	0.2595	0.0029	7.2076
		OLS	0.0972	0.0016	2.5404	0.9368	0.0139	23.5542
	a_2	FM	0.1393	0.0402	3.8793	0.2693	0.0927	7.6158
		OLS	0.1154	0.1071	2.9104	1.4685	1.3594	34.7798
1.0	a_1	FM	0.1752	0.0034	4.7026	0.2283	0.0023	6.4031
		OLS	0.1179	0.0000	3.0795	0.9316	0.0005	23.3743
	a_2	FM	0.1462	0.0642	4.0367	0.2443	0.1214	7.0874
		OLS	0.1785	0.1763	4.2729	1.8731	1.8489	42.0375

(Continued)

Table 3 (Continued)

			$T = 100$					
			Coefficients			t-statistics		
μ			MAD	Bias_{avg}	RMSE_{avg}	MAD	Biasa_{vg}	RMSE_{avg}
0.0	a_1	FM	0.1004	-0.0014	1.8474	0.2139	-0.0028	3.9817
		OLS	0.0518	-0.0011	0.9392	0.9495	-0.0187	16.8744
	a_2	FM	0.0693	-0.0001	1.3332	0.2106	0.0010	4.0022
		OLS	0.0322	-0.0004	0.6029	0.9451	-0.0082	16.7971
0.4	a_1	FM	0.1009	0.0011	1.8665	0.1984	0.0015	3.7250
		OLS	0.0564	-0.0002	1.0269	0.9610	-0.0024	17.0804
	a_2	FM	0.0704	0.0074	1.3535	0.1989	0.0240	3.7992
		OLS	0.0462	0.0380	0.8488	1.2563	1.0274	21.6872
0.6	a_1	FM	0.1025	0.0013	1.8880	0.1734	0.0014	3.2581
		OLS	0.0669	0.0009	1.2180	0.9521	0.0108	16.8886
	a_2	FM	0.0715	0.0158	1.3875	0.1732	0.0415	3.3367
		OLS	0.0789	0.0771	1.3662	1.7791	1.7358	28.7025
1.0	a_1	FM	0.1031	0.0007	1.9057	0.1580	0.0000	2.9825
		OLS	0.0739	-0.0002	1.3455	0.9570	-0.0003	16.9988
	a_2	FM	0.0726	0.0201	1.4039	0.1590	0.0499	3.0777
		OLS	0.0975	0.0968	1.6385	1.9998	1.9853	31.3980

Table 4 Summary Statistics ($d = 4$, $T = 200$ and $T = 500$)

			$T = 200$					
			Coefficients			t-statistics		
μ			MAD	Bias_{avg}	RMSE_{avg}	MAD	Bias_{avg}	RMSE_{avg}
0.0	a_1	FM	0.0628	0.0008	0.8093	0.1599	0.0014	2.0620
		OLS	0.0338	0.0003	0.4293	0.9642	0.0113	12.0811
	a_2	FM	0.0353	0.0000	0.4883	0.1577	0.0009	2.0837
		OLS	0.0163	0.0000	0.2160	0.9541	0.0049	11.9981
0.4	a_1	FM	0.0633	0.0004	0.8177	0.1496	0.0011	1.9285
		OLS	0.0367	-0.0003	0.4675	0.9697	-0.0056	12.1685
	a_2	FM	0.0350	0.0019	0.4842	0.1457	0.0078	1.9170
		OLS	0.0237	0.0199	0.3096	1.2800	1.0605	15.6250
0.6	a_1	FM	0.0624	0.0008	0.8041	0.1374	0.0019	1.7694
		OLS	0.0395	0.0001	0.5042	0.9661	0.0021	12.1435
	a_2	FM	0.0355	0.0029	0.4935	0.1348	0.0124	1.7920
		OLS	0.0317	0.0300	0.3981	1.5765	1.4835	18.5050
1.0	a_1	FM	0.0639	0.0003	0.8228	0.1166	-0.0004	1.5016
		OLS	0.0485	0.0010	0.6147	0.9781	0.0156	12.2112
	a_2	FM	0.0359	0.0052	0.4971	0.1140	0.0176	1.5087
		OLS	0.0508	0.0505	0.6036	2.0673	2.0529	22.8058

(Continued)

Table 4 (Continued)

			$T = 500$					
			Coefficients			t-statistics		
μ			MAD	Bias_{avg}	RMSE_{avg}	MAD	Bias_{avg}	RMSE_{avg}
0.0	a_1	FM	0.0347	0.0002	0.2790	0.1071	0.0012	0.8586
		OLS	0.0199	0.0000	0.1589	0.9632	0.0013	7.6666
	a_2	FM	0.0140	0.0001	0.1246	0.1042	0.0005	0.8576
		OLS	0.0065	0.0000	0.0545	0.9577	0.0026	7.6285
0.4	a_1	FM	0.0352	0.0001	0.2832	0.1005	0.0001	0.8058
		OLS	0.0216	0.0004	0.1723	0.9697	0.0161	7.7012
	a_2	FM	0.0139	0.0001	0.1218	0.0966	0.0014	0.7933
		OLS	0.0096	0.0081	0.0784	1.3030	1.0949	10.0199
0.6	a_1	FM	0.0350	0.0001	0.2817	0.0919	0.0003	0.7359
		OLS	0.0236	0.0001	0.1876	0.9773	0.0052	7.7334
	a_2	FM	0.0138	0.0006	0.1225	0.0876	0.0044	0.7228
		OLS	0.0129	0.0123	0.1024	1.6115	1.5318	11.8921
1.0	a_1	FM	0.0355	0.0004	0.2844	0.0771	0.0006	0.6157
		OLS	0.0287	0.0000	0.2291	0.9829	0.0014	7.7971
	a_2	FM	0.0140	0.0008	0.1238	0.0735	0.0046	0.6068
		OLS	0.0207	0.0206	0.1552	2.1131	2.1014	14.7066

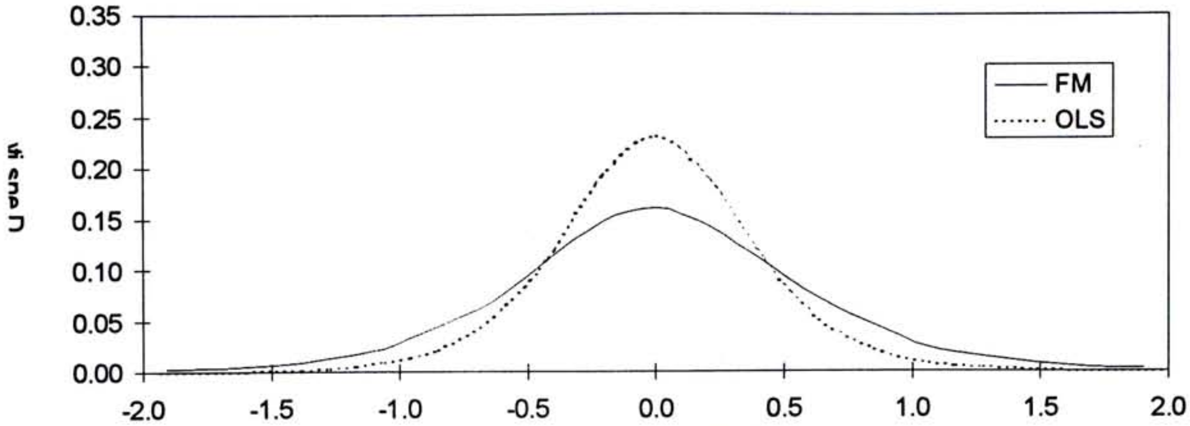

Figure 1 (d) Distributions of $T^{1/2}(\widehat{a}_1^+ - a_1)$ and $T^{1/2}(\widehat{a}_1 - a_1)$ when $T = 50$, $\mu = 1$ and $d = 4$.

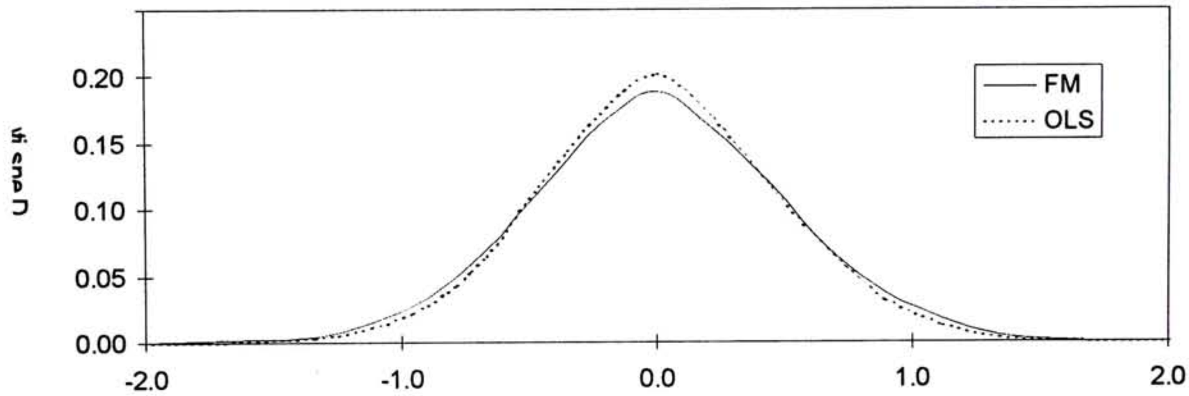

Figure 1 (e) Distributions of $T^{1/2}(\widehat{a}_1^+ - a_1)$ and $T^{1/2}(\widehat{a}_1 - a_1)$ when $T = 500$, $\mu = 0$ and $d = 1$.

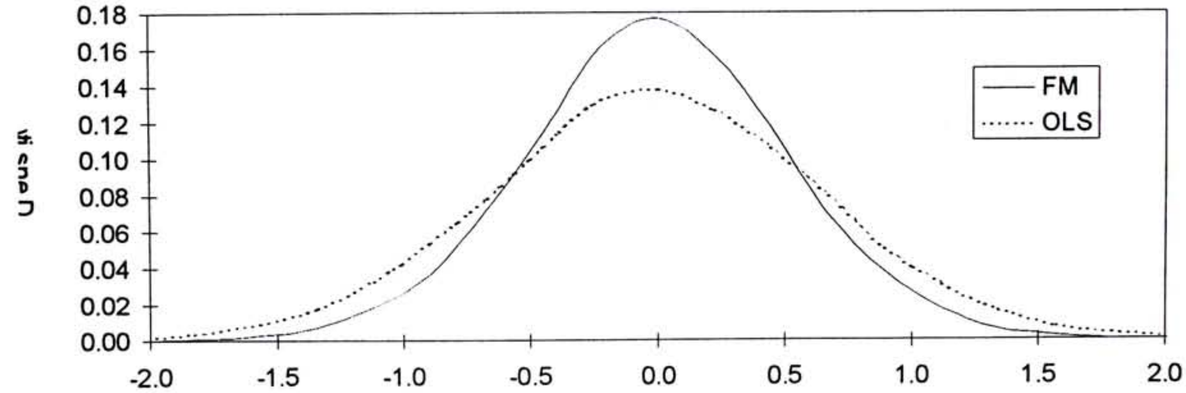

Figure 1 (f) Distributions of $T^{1/2}(\widehat{a}_1^+ - a_1)$ and $T^{1/2}(\widehat{a}_1 - a_1)$ when $T = 500$, $\mu = 1$ and $d = 1$.

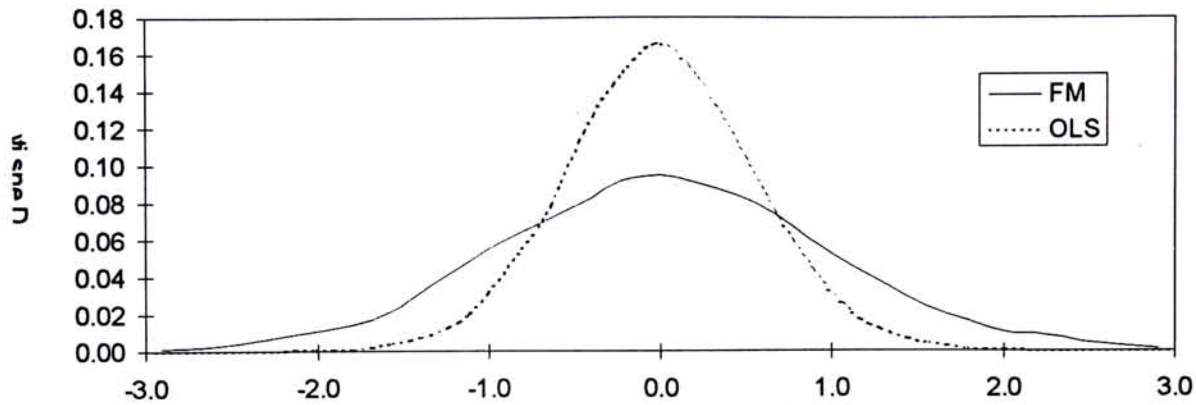

Figure 1 (g) Distributions of $T^{1/2}(\widehat{a}_1^+ - a_1)$ and $T^{1/2}(\widehat{a}_1 - a_1)$ when $T = 500$, $\mu = 0$ and $d = 4$.

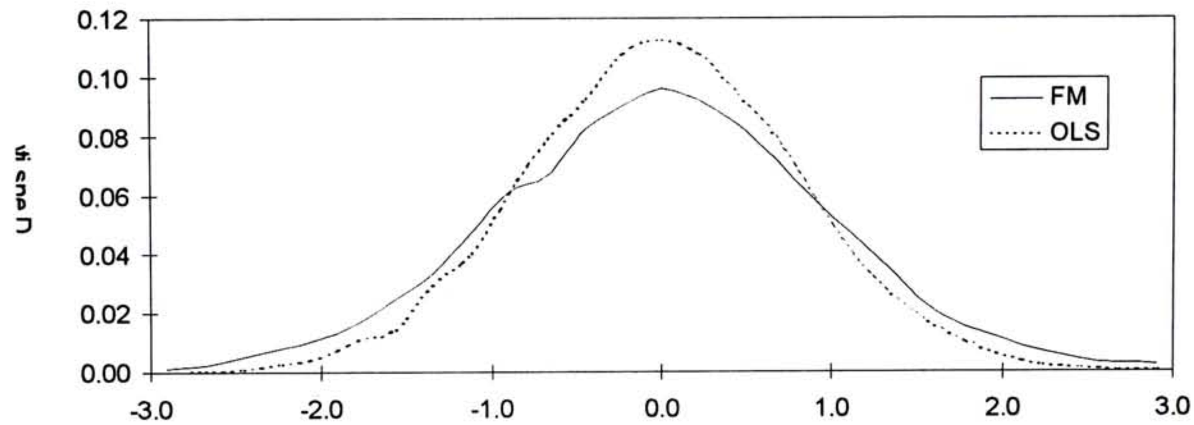

Figure 1 (h) Distributions of $T^{1/2}(\widehat{a}_1^+ - a_1)$ and $T^{1/2}(\widehat{a}_1 - a_1)$ when $T = 500$, $\mu = 1$ and $d = 4$.

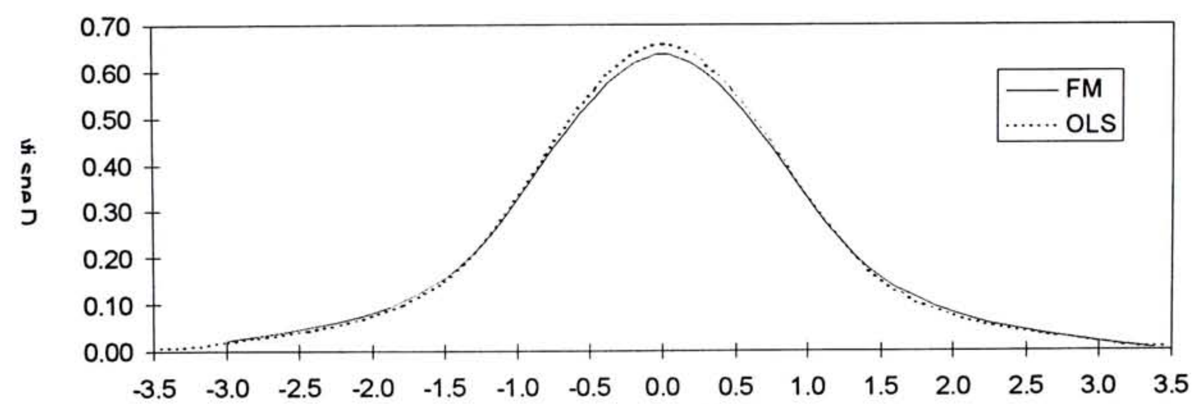

Figure 2 (a) Distributions of $T(\widehat{a}_2^+ - a_2)$ and $T(\widehat{a}_2 - a_2)$ when $T = 50$, $\mu = 0$ and $d = 1$.

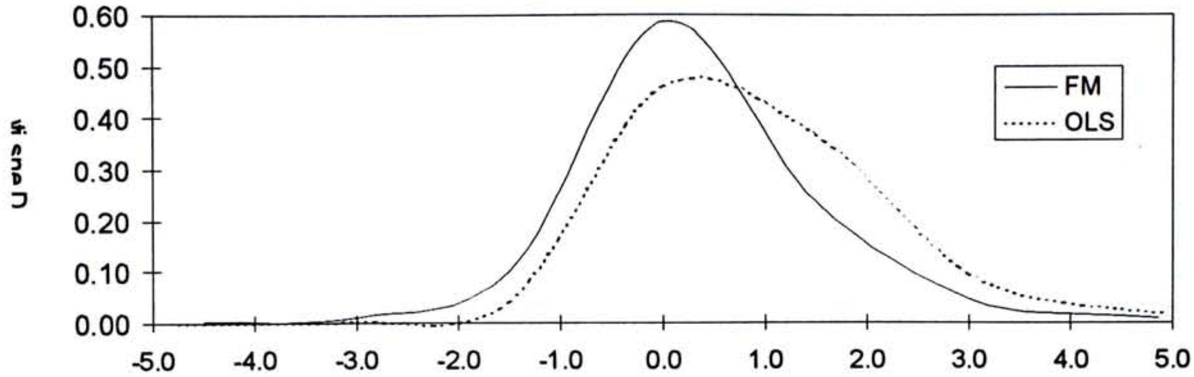

Figure 2 (b) Distributions of $T(\widehat{a}_2^+ - a_2)$ and $T(\widehat{a}_2 - a_2)$ when $T = 50$, $\mu = 1$ and $d = 1$.

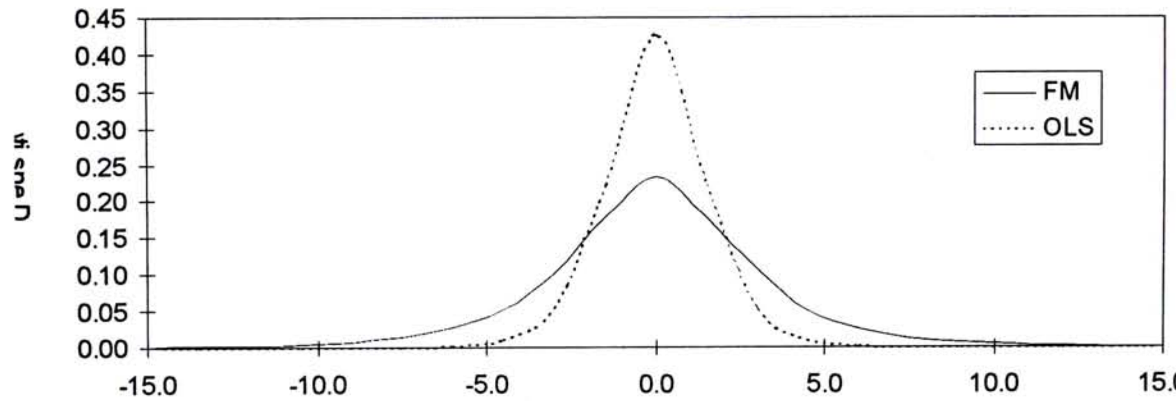

Figure 2 (c) Distributions of $T(\widehat{a}_2^+ - a_2)$ and $T(\widehat{a}_2 - a_2)$ when $T = 50$, $\mu = 0$ and $d = 4$.

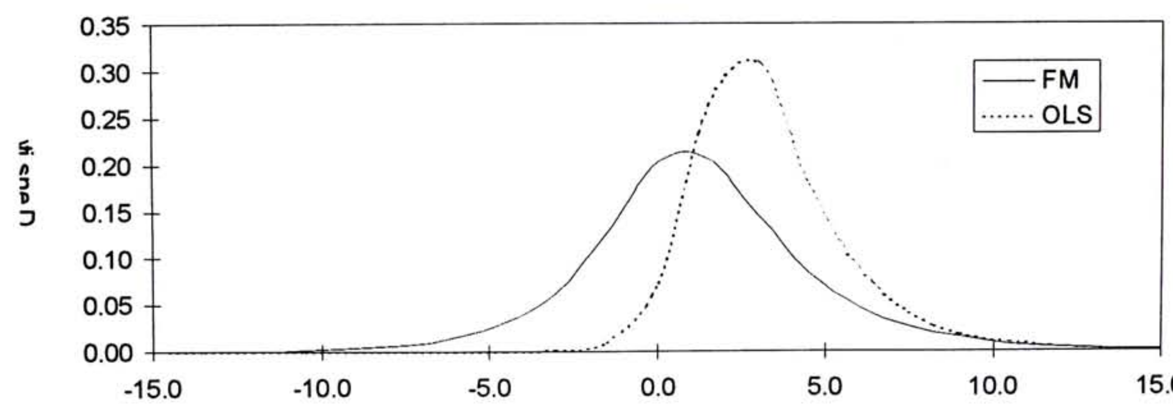

Figure 2 (d) Distributions of $T(\widehat{a}_2^+ - a_2)$ and $T(\widehat{a}_2 - a_2)$ when $T = 50$, $\mu = 1$ and $d = 4$.

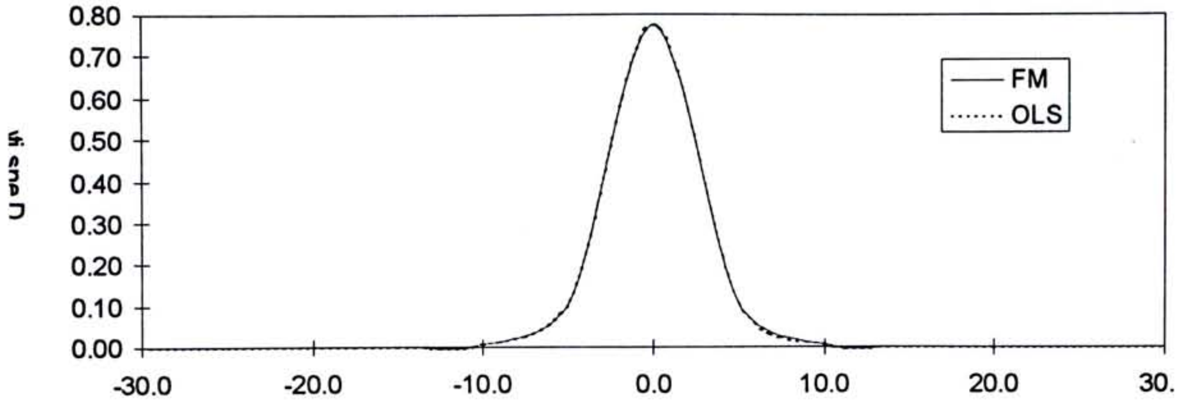

Figure 2 (e) Distributions of $T(\widehat{a}_2^+ - a_2)$ and $T(\widehat{a}_2 - a_2)$ when $T = 500$, $\mu = 0$ and $d = 1$.

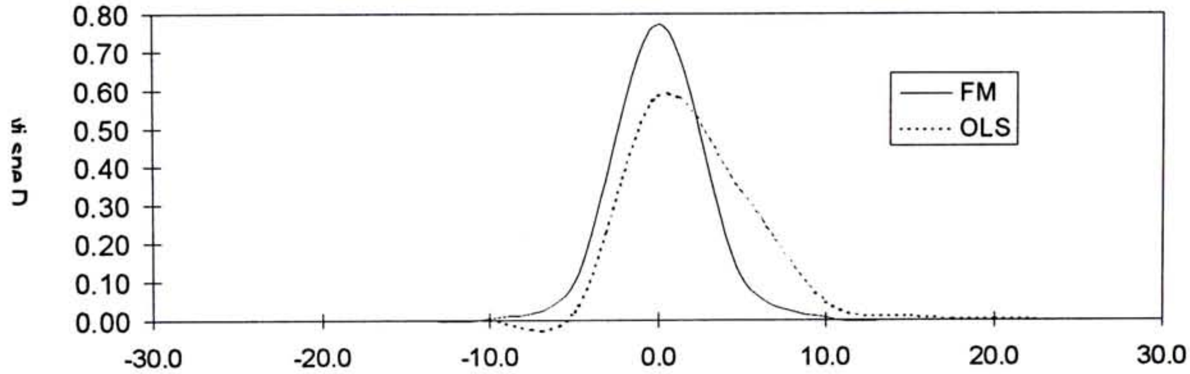

Figure 2 (f) Distributions of $T(\widehat{a}_2^+ - a_2)$ and $T(\widehat{a}_2 - a_2)$ when $T = 500$, $\mu = 1$ and $d = 1$.

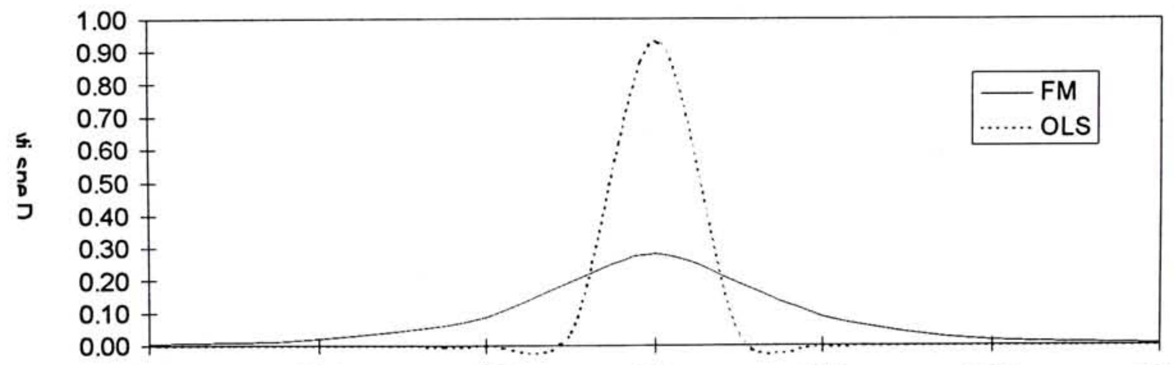

Figure 2 (g) Distributions of $T(\widehat{a}_2^+ - a_2)$ and $T(\widehat{a}_2 - a_2)$ when $T = 500$, $\mu = 0$ and $d = 4$.

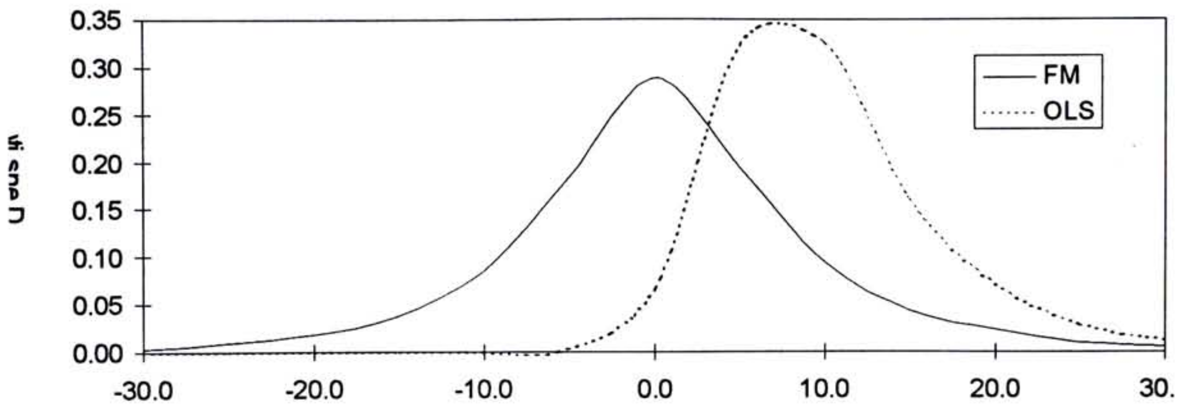

Figure 2 (h) Distributions of $T(\widehat{a}_2^+ - a_2)$ and $T(\widehat{a}_2 - a_2)$ when $T = 500$, $\mu = 1$ and $d = 1$.

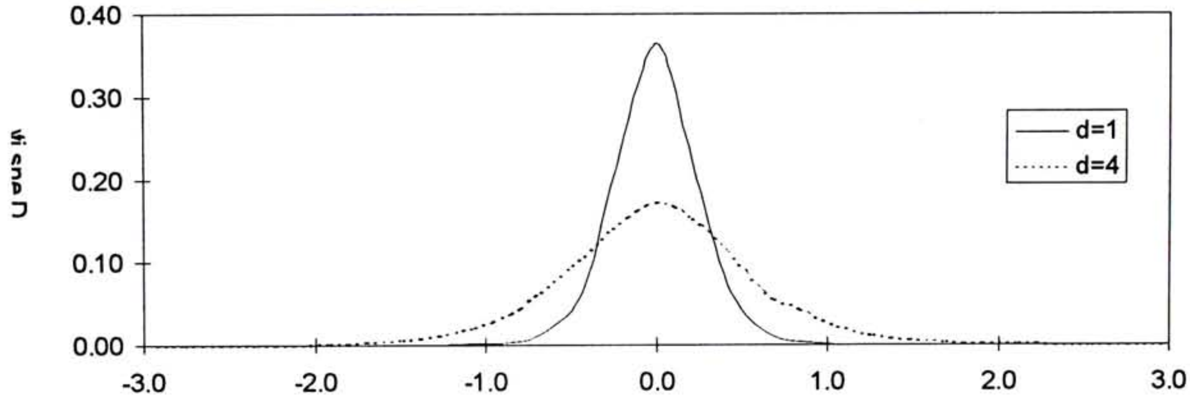

Figure 3 (a) Comparison of distributions of $T^{1/2}(\widehat{a}_1^+ - a_1)$ when $T = 50$ and $\mu = 0$.

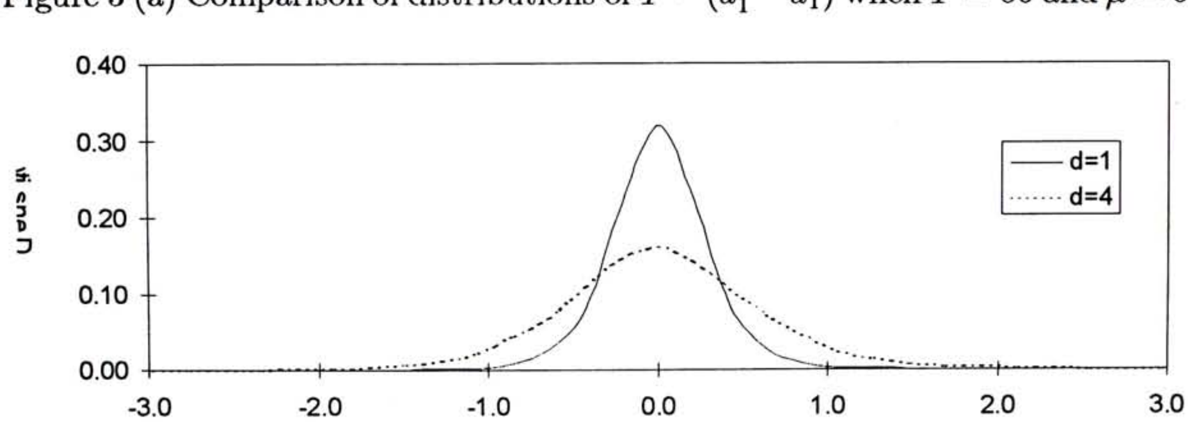

Figure 3 (b) Comparison of distributions of $T^{1/2}(\widehat{a}_1^+ - a_1)$ when $T = 50$ and $\mu = 1$.

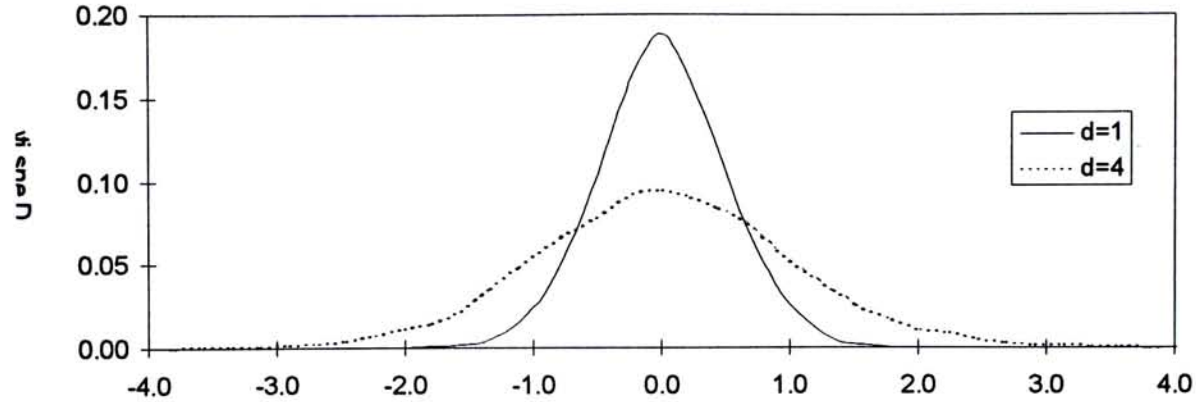

Figure 3 (c) Comparison of distributions of $T^{1/2}(\widehat{a}_1^+ - a_1)$ when $T = 500$ and $\mu = 0$.

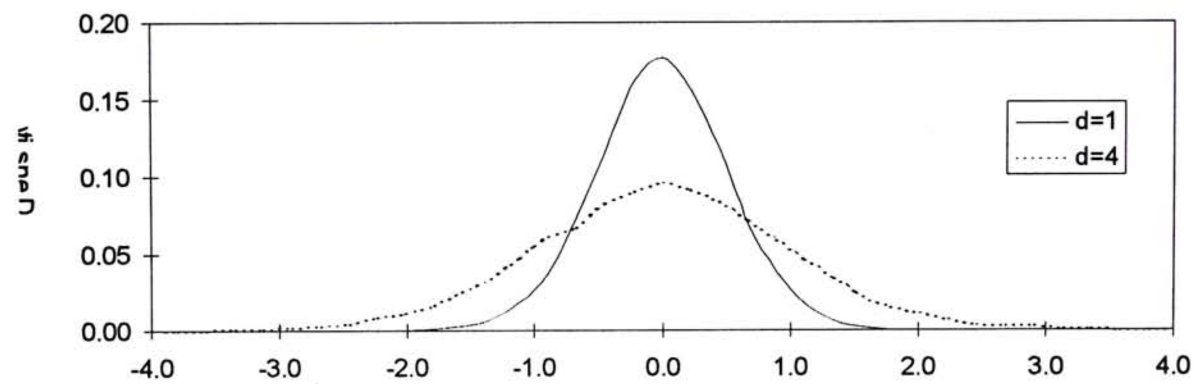

Figure 3 (d) Comparison of distributions of $T^{1/2}(\widehat{a}_1^+ - a_1)$ when $T = 500$ and $\mu = 1$.

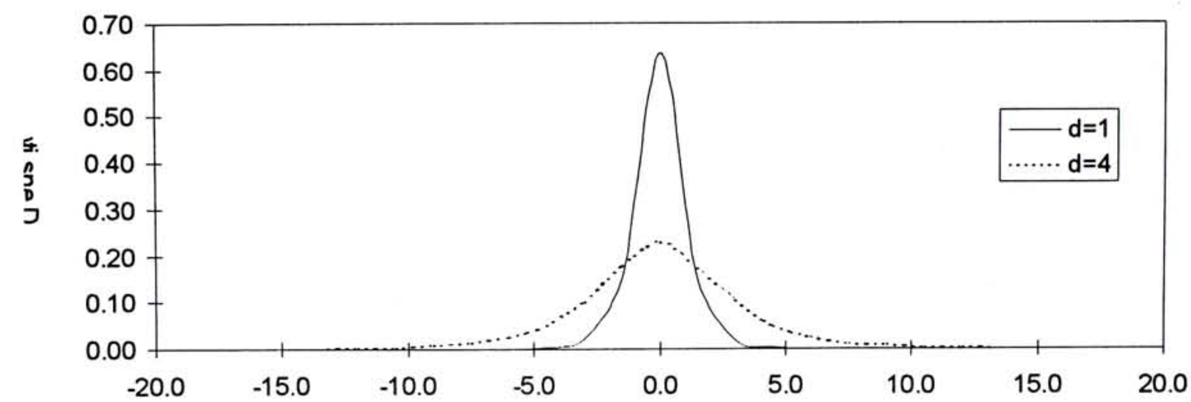

Figure 4 (a) Comparison of distributions of $T(\widehat{a}_2^+ - a_2)$ when $T = 50$ and $\mu = 0$.

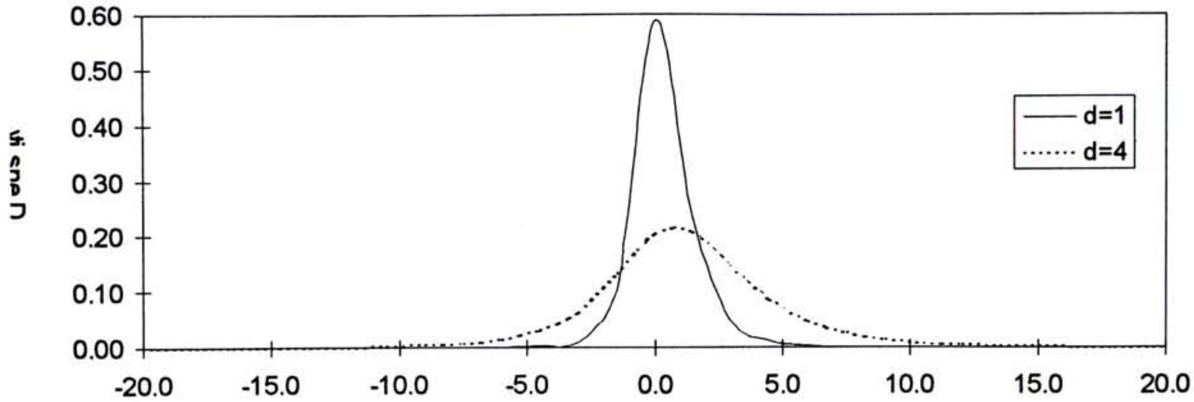

Figure 4 (b) Comparison of distributions of $T(\widehat{a}_2^+ - a_2)$ when $T = 50$ and $\mu = 1$.

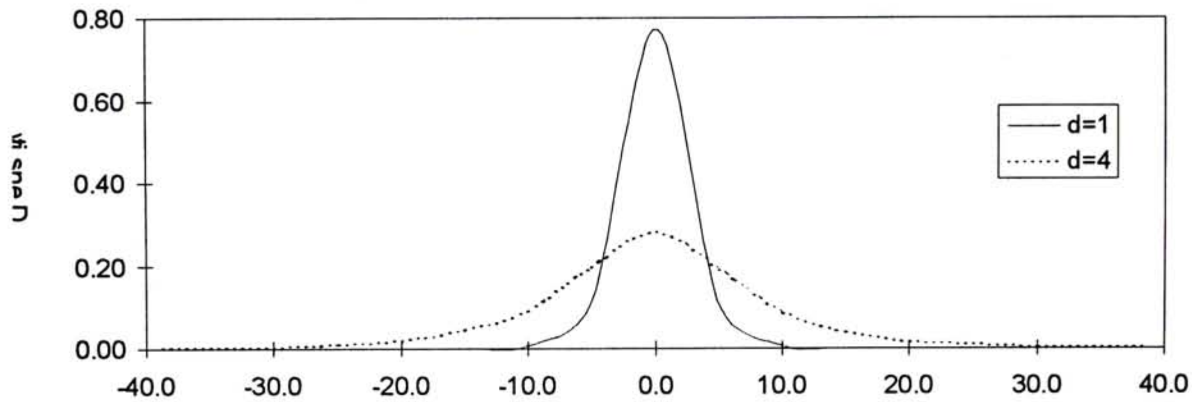

Figure 4 (c) Comparison of distributions of $T(\widehat{a}_2^+ - a_2)$ when $T = 500$ and $\mu = 0$.

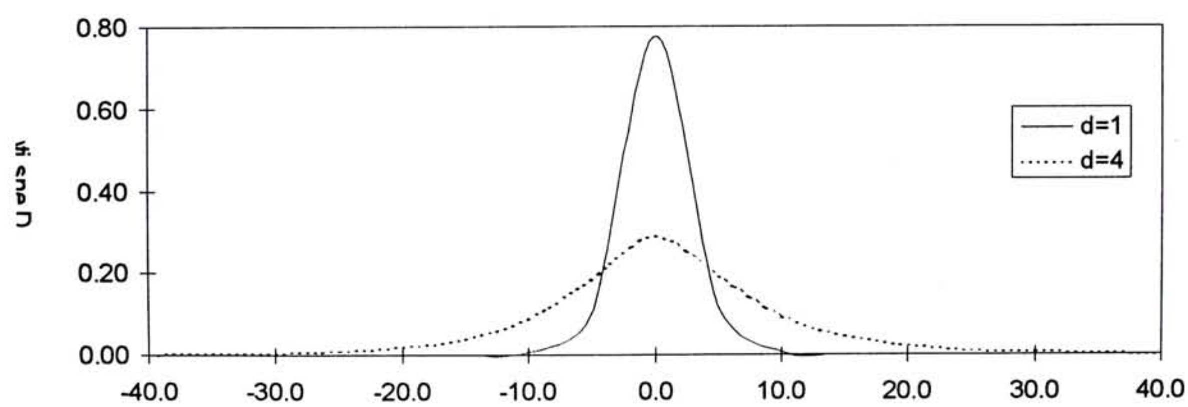

Figure 4 (d) Comparison of distributions of $T(\widehat{a}_2^+ - a_2)$ when $T = 50$ and $\mu = 0$.

8. References

ANDREWS, D. W. K. (1991): "Heteroskedasticity and Autocorrelation Consistent Covariance Matrix Estimation," *Econometrica*, 59, 817-858.

BILLINGSLEY, P. (1968): *Convergence of Probability Measures*, New York : Wiley.

CANOVA, F. AND B. E. HANSEN (1995): "Are Seasonal Patterns Constant Over Time ? A Test for Seasonal Stability," *Journal of Business and Economic Statistics*, 13, 237-252.

CHAN, N. H. (1989): "Asymptotic Inference for Unstable Autoregressive Time Series with Drifts," *Journal of Statistical Planning and Inference*, 23, 301-312.

——— AND C. Z. Wei (1988): "Limiting Distributions of Least Squares Estimates of Unstable Autoregressive Processes," *Annals of Statistics*, 16, 367-401.

DAVIDSON, J. E. H., D. F. HENDRY, F. SRBA AND S. YEO (1978): "Econometric Modelling of the Aggregate Time-Series Relationship Between Consumers' Expenditure and Income in the United Kingdom," *Economic Journal*, 88, 661-692.

DICKEY, D. A. AND W. A. FULLER (1979): "Distribution of the Estimators for Autoregressive Time Series With a Unit Root," *Journal of the American Statistical Association*, 74, 427-431.

———, D. P. HASZA AND W.A. FULLER (1984): "Testing for Unit Roots in Seasonal Time Series," *Journal of the American Statistical Association*, 79, 355-367.

ENGLE, R. F. AND C. W. J. GRANGER (1987): "Cointegration and Error-Correction : Representation, Estimation and Testing," *Econometrica*, 55, 251-276.

FULLER, W. A. (1976): *Introduction to Statistical Time Series*, New York : Wiley.

HANNAN, E. J. (1970): *Multiple Time Series*, New York : Wiley.

HANSEN, B. E. (1992): "Efficient Estimation and Testing of Cointegrating Vectors in the Presence of Deterministic Trends," *Journal of Econometrics*, 53, 87-121.

HASZA, D. P. AND W. A. FULLER (1982): "Testing for Nonstationary Parameter Specifications in Seasonal Time Series Models," *Annals of Statistics*, 10, 1209-1216.

HYLLEBERG S., R. F. ENGLE, C. W. J. GRANGER, B. S. YOO (1990): "Seasonal Integration and Cointegration," *Journal of Econometrics*, 44, 215-238.

JOHANSEN, S. (1988): "Statistical Analysis of Cointegration Vectors," *Journal of Economic Dynamics and Control*, 12, 231-254.

——— (1991): "Estimation and Hypothesis Testing of Cointegration Vectors in Gaussian Vector Autoregressive Models," *Econometrica*, 59, 1551-1580.

KAC, M. (1980): *Integration in Function Spaces and Some of Its Applications*, Pisa : Academia Nazionale Dei Lincei.

KITAMURA, Y. AND P. C. B. PHILLIPS (1995): "Efficient IV Estimation in Nonstationary Regression : An Overview and Simulation Study," *Econometric Theory*, 11, 1095-1130.

LAI, T. L. AND C. Z. WEI (1983): "Asymptotic Properties of General Autoregressive Models and Strong Consistency of Least-Squares Estimates of Their Parameters," *Journal of Multivariate Analysis*, 13, 1-23.

LEE, H. S. (1992): "Maximum Likelihood Inference on Cointegration and Seasonal Cointegration," *Journal of Econometrics*, 54, 1-47.

MAGNUS J. R. AND H. NEUDECKER (1988): *Matrix differential Calculus with Applications in Statistics and Econometrics*, New York : Wiley.

OSBORN, D. R., A. P.L. CHUI, J. P. SMITH AND C. R. BIRCHENHALL (1988): "Seasonality and the Order of Integration for Consumption," *Oxford Bulletin of Economics and Statistics*, 50, 361-377.

PARK, J. Y. (1992): "Canonical Cointegrating Regressions," *Econometrica*, 60, 119-143.

——— AND P. C. B. PHILLIPS (1988): "Statistical Inference in Regression with Integrated Process : Part 1," *Econometric Theory*, 4, 468-497.

PHILLIPS, P. C. B. (1986): "Understanding Spurious Regressions in Econometrics," *Journal of Econometrics*, 33, 311-340.

——— (1987): "Time Series Regression with a Unit Root," *Econometrica*, 55, 277-301.

——— (1988a): "Multiple Regression with Integrated Processes," in N.U. Prabhu (ed.), *Statistical Inference from Statistical Processes, Contemporary Mathematics*, 80, 79-106.

——— (1988b): "Weak Convergence to the Matrix Stochastic Integral $\int_0^1 BdB'$," *Journal of Multivariate Analysis*, 24, 252-264.

——— (1991a): "Optimal Inference in Cointegrated Systems," *Econometrica*, 59, 238-306.

——— (1991b): "Spectral Regression for Cointegrated Time Series," in W. Barnett, J. Powell and G. Tauchen (eds.), *Nonparametric and Semiparametric Methods in Economics and Statistics*, New York : Cambridge University Press.

——— (1993): "Fully Modified Least Squares and Vector Autoregression," Cowles Foundation Discussion Paper, No. 1047.

——— (1995): "Fully Modified Least Squares and Vector Autoregression," *Econometrica*, 63, 1023-1078.

——— AND S. N. DURLAUF (1986): "Multiple Time Series with Integrated Variables," *Review of Economic Studies*, 53, 473-496.

——— AND B. E. HANSEN (1990): "Statistical Inference in Instrumental Variables Regressions with I(1) Processes," *Review of Economic Studies*, 57, 99-125.

——— AND J. Y. PARK (1988): "Asymptotic Equivalence of Ordinary Least Squares and Generalized Least Squares in Regressions and Integrated Regressors," *Journal of the American Statistical Association*, 83, 111-115.

——— AND V. SOLO (1992): "Asymptotics for Linear Processes," *Annals of Statistics*, 20, 971-1001.

POLLARD, D. (1984): *Convergence of Stochastic Processes*, New York : Springer-Verlag.

PRIESTLEY, M. B. (1981): *Spectral Analysis and Time Series, Volumes I and II*, New York : Academic Press.

RAO, M. M. (1978): "Asymptotic Distribution of an Estimator of the Boundary Parameter of an Unstable Process," *Annals of Statistics*, 6, 185-190.

SIMS, C. A., J. H. STOCK AND M. W. WATSON (1990): "Inference in Linear Time Series Models with Some Unit Roots," *Econometrica*, 58, 113-144.

STOCK, J. H. (1987): "Asymptotic Properties of Least Squares Estimators of Cointegrating Vectors," *Econometrica*, 55, 1035-1056.

TIAO, G. C. AND R. S. TSAY (1983): "Consistency Properties of Least Squares Estimates of Autoregressive Parameters in ARMA Models," *Annals of Statistics*, 11, 856-871.

TSAY, R. S. AND G. C. TIAO (1990): "Asymptotic Properties of Multivariate Nonstationary Processes with Applications," *Annals of Statistics*, 18, 220-250.

WEST, K. D. (1988): "Asymptotic Normality When Regressors Have a Unit Root," *Econometrica*, 56, 1397-1417.